Online Social Signals and Consumer Behavior

Qianran(Jenny) Jin
Desautels Faculty of Management
McGill University
Montreal, Quebec

June 2020

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

TABLE OF CONTENTS

TABLE OF TABLES

TABLE OF FIGURES

Abstract

With the exponential increase in the amount of information generated nowadays, online social signals have become an important source of information for consumers to reduce the uncertainties in their decision making. The core of this thesis comprises three essays that investigate the nature and effects of online social signals. In the first essay, we propose a conceptual framework that integrates and extends prior studies on online social signals. The goal of having this framework is to systematically understand the different types of online social signals and to examine the differential impact of these online social signals on consumer decision making. The second and third essays are empirical studies based on the research gaps that we have identified in essay one. In the second essay, we examine the relative strength of one online social signal, popularity information, on individuals' decisions in the presence of position effect. We specifically examine the effect of popularity on the consumer's two-stage decision process. Our results show that the effect of popularity could be potentially overestimated if not accounting for the position, and that popularity has a differential impact on the consumer's search and choice decisions. In the third essay, we study how consumers make decisions in the presence of two types of online social signals: pre-consumption signals and post-consumption signals. We find that the influence of these two types of signals depends on how the choice sets are constructed. While the post-consumption signal remains the stronger signal for consumers' decisions, consumers will only use this signal when they are cognitively capable. Otherwise, they will rely on the pre-consumption signal. We also show that consumers will change their decisions when the post-consumption signal is presented to them explicitly. This dissertation makes contributions to the burgeoning stream of literature by proposing an overarching framework on online social signals that lays the foundation for further exploration. It also

contributes to the existing literature with practical implications on how platforms can display these online social signals and how policymakers can utilize online social signals for designing policy.

Résumé

Avec l'augmentation exponentielle de la quantité d'informations générées de nos jours, les signaux sociaux en ligne sont devenus une source importante d'informations pour les consommateurs afin de réduire les incertitudes dans leur prise de décision. Le cœur de cette thèse comprend trois essais qui étudient la nature et les effets des signaux sociaux en ligne. Dans le premier essai, nous proposons un cadre conceptuel qui intègre et prolonge les études antérieures sur les signaux sociaux en ligne. L'objectif de ce cadre est de comprendre systématiquement les différents types de signaux sociaux en ligne et d'examiner l'impact différentiel de ces signaux sociaux en ligne sur la prise de décision des consommateurs. Les deuxième et troisième essais sont des études empiriques basées sur les lacunes de recherche que nous avons identifiées dans le premier essai. Dans le deuxième essai, nous examinons la force relative d'un signal social en ligne, les informations de popularité, sur les décisions des individus en présence d'un effet de position. Nous examinons spécifiquement l’effet de la popularité sur le processus de décision en deux étapes du consommateur. Nos résultats montrent que l’effet de la popularité pourrait être surestimé s’il ne tient pas compte du poste, et que la popularité a un impact différent sur les décisions de recherche et de choix du consommateur. Dans le troisième essai, nous étudions comment les consommateurs prennent des décisions en présence de deux types de signaux sociaux en ligne: les signaux de pré-consommation et les signaux de post-consommation. Nous constatons que l'influence de ces deux types de signaux dépend de la manière dont les ensembles de choix sont construits. Alors que le signal de post-consommation reste le signal le plus fort pour les décisions des consommateurs, les consommateurs n'utiliseront ce signal que lorsqu'ils seront capables cognitivement. Sinon, ils s'appuieront sur le signal de pré-consommation. Nous montrons également que les consommateurs changeront leurs décisions lorsque le signal post-

consommation leur sera présenté explicitement. Cette thèse contribue au flux de littérature en plein essor en proposant un cadre global sur les signaux sociaux en ligne qui jette les bases d'une exploration plus approfondie. Il contribue également à la littérature existante avec des implications pratiques sur la manière dont les plateformes peuvent afficher ces signaux sociaux en ligne et sur la manière dont les décideurs peuvent utiliser les signaux sociaux en ligne pour concevoir des politiques.

Acknowledgments

First and foremost, I would like to express my deepest appreciation to my advisors, Alain Pinsonneault and Animesh, for their guidance throughout my doctoral studies. This dissertation could not have been completed without both of their mentorship, support, and patience. I greatly appreciate Animesh for the tremendous amount of time and effort he spent with me. Whenever and wherever I had ideas, questions, or issues, Animesh always spared the time to discuss them with me and brought a lot of insights to the discussion. I am also truly grateful to Alain for his care and insightful guidance. Alain has provided not only incredible suggestions on my research but also invaluable recommendations on building my career as an academic. I truly feel lucky to have two advisors who greatly supported me throughout my doctoral studies.

I would also like to express my thanks to Jui Ramaprasad for her remarkable support. Jui has been a role model for me as a female academic and has always lent an ear as a friend to listen to my personal problems.

I would like to thank my dissertation committee members Nathan Yang and SunAh Kim, for their assistance and advice over the course of my dissertation work. I would also like to extend my sincere thanks to all the faculty members in the IS Department, Kartik Ganju, Genevieve Bassellier, Liette Lapointe, Emma Vaast, Warut Khernamnuai, Kunsoo Han, and Changseung Yoo (ordered by the distance from my office to their offices), for giving me valuable advice and relentless support. And thank you, Rola Zoayter, for the kind admin support. I cannot leave Desautels without mentioning the support from the Ph.D. office, Samer Faraj, Stella Scalia, and Yanyan Shu. Thank you for always being there for the Ph.D. students. Thanks also to my friends and colleagues whom we have spent days and nights together on the

Bronfman 5th floor. Thanks to June, Takumi, Divinius, Ghazelah, Karla, Zhu, Changjie, Weiyu, Yonghwan, Kyunghee, Hyunji, and Yinan.

This journey would not have been possible without the support of my family. Thank you to my parents, Lili and Xiaofeng, who have set good examples to me as academics. I still remember the days my mother had to take care of me and wrote her own doctoral thesis when I was in elementary school, and my father always went to his office and worked till late at night. I remember waking up at 2 am and finding my parents were still working. Although my parents have been very stringent, they have always been very liberal and supportive of my decisions. Thank you to my grandparents for believing in me and encouraging me to grow up as an independent female. Thank you to my husband, *Quan*, who has been extremely patient and unwaveringly supportive. I use Italic instead of the agreed 12.5 font on your name to acknowledge your incredible support in helping me finish my thesis

Contribution of Authors

All three essays are co-authored by Qianran (Jenny) Jin, the first author, and Animesh Animesh and Alain Pinsonneault, as the second and third authors, respectively. However, the first author performed the vast majority of the work, while the second and third authors provided important guidance and feedback to the first author on how to improve the paper. Earlier versions of the dissertation were presented at the International Conference on Information Systems (ICIS) 2019 (Essay 2, Essay 3), Conference on Information Systems and Technology (CIST) 2019 (Essay 2), and Statistical Challenges in E-Commerce Research (SCECR) 2019 (Essay 2).

Chapter 1. Introduction

In the Internet age, the amount of information and data that are available to us has continued to grow. Each day, there are 2.5 quintillion bytes of data created (Marr 2018). While the fast-growing information has allowed us to access more information, it has resulted in the availability of a vast number of alternatives. For example, Amazon sells more than 12 million products (Dayton 2020), Spotify has over 50 million songs on their platform (Spotify 2020), and YouTube has more than 720,000 hours of video uploaded to the platform every day (Mohsin 2020). These large numbers of options have made it difficult for consumers to discover, filter, compare and select an alternative to consume. To make it easier for consumers to manage the information overload issue and to process a large amount of information, online platforms start to provide information signals that can be categorized as **online social signals**. For example, online shopping website displays the sales information, which is an aggregation of how many prior consumers have purchased the product. Online video platforms display the percentage of Thumbs Up videos have. Different from the other descriptive information that is displayed online, online social signals can help consumers reduce the uncertainties associated with the products and help them make better decisions.

As online social signals have become an important factor for consumer's decision making, it is vital for both researchers and practitioners to understand the role of different types of online social signals on consumer behavior. While prior studies have examined online social signals separately in multiple contexts, different labels have been used to refer to the same signal, and the same label has been used to refer different signals (Adomavicius et al. 2013, Cai et al. 2009, Chevalier and Mayzlin 2006, Chintagunta et al. 2010, Dewan and Ramaprasad 2012, Goes et al. 2014, Lee et al. 2015, Mudambi and Schuff 2010, Salganik et al. 2006, Tucker and

Zhang 2011). Therefore, it is challenging to synthesize and compare previous findings. Motivated by the growth of different types of online social signals and the need for a comprehensive framework in understanding the impact of online social signals on consumer behavior, we propose three essays in this dissertation to fill in the gap.

The first essay addresses the gap by developing a conceptual framework for online social signals. We first discuss the process of how online social signals are generated. Based on the generation process, we propose a taxonomy of online social signals with two dimensions: information type and information source. We explore different subtypes of online social signals with this taxonomy. We try to understand how and why these online social signals can have an influence on consumer behavior. Finally, we propose the conceptual framework of how online social signals can impact prospective consumer's decision making by explicitly examining the search, consumption, and post-consumption decisions. We also discuss two sources of heterogeneity: product characteristics and consumer characteristics. This study allows us to systematically synthesize the results from prior studies and provides a road map for both future researchers and practitioners.

Drawing on the conceptual framework developed in the first essay, the second essay and the third essay empirically examine two research gaps we identify from the framework in the context of the digital good. In the second essay, we investigate the impact of one of the online social signals, popularity, on consumer's search decision and choice decision. Prior studies have shown that popularity information has a great impact on an individual's decision (Dewan et al. 2017, Salganik et al. 2006, Tucker and Zhang 2011, Zhang 2010). However, they have largely overlooked the influence of position in their studies, which results from our top-to-bottom reading habit. Without accounting for the position effect, we might be overestimating the popularity effect. In addition, it is important and necessary to consider different decision stages because a different amount of information is taken into account for each decision. The search decision allows consumers to obtain information about the product and better construct their preferences when they do not know much about the product. Therefore, they are likely to be influenced by the popularity information. After consumers have made the search decision, they have obtained some information. Thus, the effect of online social signals might be weaker. However, prior studies have only examined the influence of popularity effect on either search stage or choice stage. Taken together, we adopt an integrated approach to compare and contrast the effect of both popularity information and position on consumer's two-stage decisions. We randomize the listing of products to control for the potential position effect, and our experiments allow us to observe participants' sequential decision making (search and choice) so that we can clearly understand the effect of popularity on consumer's two-stage decision making.

In the third essay, we examine how consumers use two types of online social signals in making decisions. Prior studies on consumer decision making among different alternatives have mostly focused on products' inherent attributes (Evangelidis and Levav 2013, Simonson 1989, Tversky and Simonson 1993). However, with the growth of technology, online platforms can provide aggregated data generated by past and current consumers to prospective consumers that act as a social signal of product quality. Among the signals, pre-consumption signals (i.e., information from prior consumption) and post-consumption signals (i.e., information from post-consumption) are the most common ones. Research has shown that both pre-consumption signals and post-consumption signals influence consumer's decision making independently (Cai et al. 2009, Chevalier and Mayzlin 2006, Chintagunta et al. 2010, Zhang and Liu 2012). However, these two signals are interrelated and should be used together to make decisions. Therefore, we first use the Bayesian Inference model to theoretically predict consumer's choices to account for these two interrelated signals. Then we conduct a series of experiments to understand how consumers make their choices by providing choice pairs in two scenarios: (1) one option is seemingly dominant over the other option, (2) two options are conflicting. By examining these two scenarios, we cover all the possible combinations of how options can be constructed with pre-consumption signals or post-consumption signals.

The next three chapters consist of these three essays. Chapter 2, entitled "*A Conceptual Framework for Online Social Signals*", introduces the conceptual framework by describing its components. Chapter 3, entitled "*Does Popularity Really Matter—Disentangling Popularity and Position Effect Using an Experimental Approach*", empirically studies the impact of popularity signal on consumer's two-stage decision process. Chapter 4, entitled "*Decision Making Under Conflicting Signals*", examines the interplay between two types of social signals. The three essays are written such that each is self-contained with its own introduction, related literature, methods, and research findings. We conclude the dissertation in Chapter 5, where we provide an overall summary and discussion of the whole thesis.

Chapter 2. A Conceptual Framework for Online Social Signals

1. Introduction

Since Dewey (1910) introduced the five-stage problem-solving process and Engel et al. (1978) extended and applied it to consumer behavior, the five-stage decision-making process has long been used to analyze how consumers make decisions. Specifically, it proposes that consumers go through the following five stages in their decision-making: *problem recognition*, *information search*, *alternative evaluation*, *consumption,* and *post-consumption evaluation*. *Problem recognition* is the first stage when consumers realize they have a need or a problem, and then they will search for the related information about this need or the problem in the *information search* stage. After that, consumers evaluate the different options they have collected in *the alternative evaluation* stage. In the fourth stage, consumers make consumption decisions based on all the alternatives they have collected in the previous steps. It then follows the *post-consumption evaluation* stage in which consumers evaluate the consumption decision they have just made.

With the development of the Internet, an increasing number of consumers have turned their decision-making process online. According to a poll conducted by NPR (2018), 69% of Americans purchase online, and 25% of Americans buy at least once a month online. As a result, the consumer's decision-making stages have also moved online. Specifically, *information search*, *alternative evaluation*, *consumption,* and *post-consumption evaluation,* these four stages have been transformed dramatically. For the *information search* stage, consumers now search online for related information about their needs. For the *alternative evaluation* stage, consumers can evaluate the different alternatives they searched in the previous stage by easily comparing the products across different dimensions using online tools or use information such as "people who viewed this product have also viewed that product" to infer the quality within choice sets. The *consumption* stage is simplified to a single click before entering payment information. While the evaluation of the consumption is offline or online depends on whether it is a physical product or digital product, the results of these evaluations are brought online. In the *post-consumption evaluation* stage, consumers can express their opinion toward the product by writing online reviews.

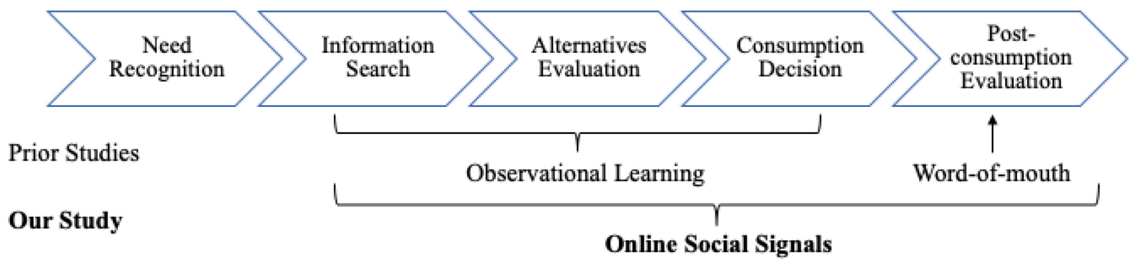
While it has become easier for an individual consumer to carry out their decision process with the aid of technology, platforms have also benefited from this digitization by getting access to consumer's digital footprint, which includes a lot of traceable digital activities and actions (Rouse 2014). These digital footprints can be classified into either a passive digital footprint, which the user did not purposefully leave the data but was unintentionally left on the platform (e.g., online cookies), or an active digital footprint, which was created by the user intentionally (e.g., online review) (Girardin et al. 2008). Many platforms have taken advantage of the digital footprint and display information that is generated from these digital footprints to help future consumers in their decision-making journey. For example, an aggregation of the digital footprints from the *information search* stage can be represented by "number of clicks". A digital footprint from the *consumption* stage can be represented by "Based on the similarity between you and your friend, we think you will also like this product". In this paper, we define *the information generated from prior consumers' digital footprint, either unintentionally or intentionally, throughout their decision stages* as **online social signals**. The word "signal" in the **online social signals** should not be confused with the economic signaling theory even though both signals are used to reduce the uncertainty related to product quality (Spence 1978).

However, the signals in the economic signaling theory are created by the sellers, whereas the online social signals in our context are created by prior consumers.

While the World Wide Web has been there for more than two decades, online social signals also have a long history and have attracted attention from academics. One stream of research uses observational learning, herding, or information cascade (Dewan and Ramaprasad 2012, Duan et al. 2009, Lee et al. 2015, Zhang and Liu 2012) to examine the information generated from *information search* stage, *alternative evaluation* stage or *consumption* stage. Another stream of research uses online word-of-mouth or user-generated content to study the information generated from the *post-consumption evaluation* stage (Chevalier and Mayzlin 2006, Chintagunta et al. 2010, Duan et al. 2008b, Lee et al. 2015). From Figure 1, we can see that these two streams of research have both studied part of the online social signals we have defined in this paper. As a result, limited studies have brought together these two streams and examined observational learning and word-of-mouth together (Chen et al. 2011, Li and Wu 2018). In order to have a better understanding of how different online social signals would impact consumers differently, there is a need for a conceptual framework that can systematically examine the effect. Without an overarching conceptual framework, it is also very hard to synthesize the existing findings and find potential research gaps. An integrated framework will also inform practitioners on how to design platforms or carry out policies. Currently, we can see different platforms are displaying different social signals. Even for the same types of platforms, they also show various kinds of social signals. Having a conceptual framework allows platforms to systematically decide what online social signals they should present to the consumers under what circumstances that would be most effective.

Note that the word "consumption" is used in a broad sense (Merriam-Webster). While consumption usually refers to the utilization of goods or services, we argue that consumption means reading, viewing, watching, listening to a digital good in the context of digital goods (such as Facebook posts, Instagram photos, YouTube videos). Therefore, our framework is not limited to purchasing products online. In addition, the word "product" is also used in a broad sense, which includes not only physical goods and services but also includes digital goods.

Figure 1 Definition of Online Social Signals in This Study

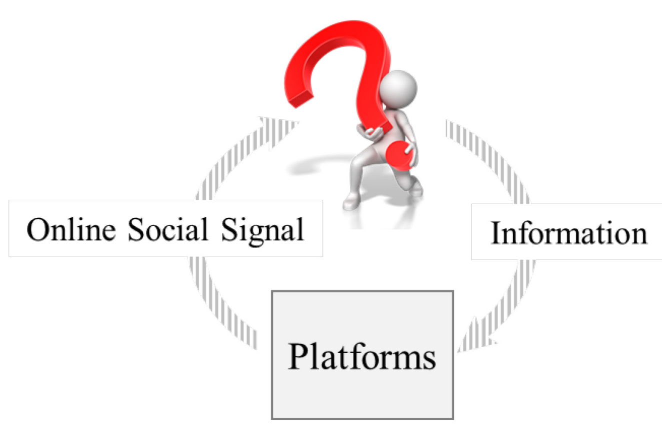


This rest of this paper proceeds as follows. Section 2 presents the process of how online social signals are generated. Section 3 describes the taxonomy of online social signals and subtypes. Section 4 explains the conceptual framework of online social signals' effect on consumer decision making. Section 5 extends the framework by introducing product characteristics and consumer characteristics. We conclude in Section 6.

2. Online Social Signals Generation Process

To understand why online social signals have different impacts on consumers, it is important to first understand how online social signals are generated. Assuming consumer A is the first consumer that goes to the platform. When consumer A lands on the platform, he/she performs a task and then leaves the platform. The information from consumer A, including consumer A's actions and activities, is recorded by the platform and is used to generate online social signals for future consumer B. When consumer B lands on the same platform, he/she sees the online social signals generated from consumer A. Consumer B leaves the platform after he/she finishes the task. The information from consumer B, together with the information from consumer A is now used to generate online social signals for future consumer C, so on and so forth. Therefore, at any point in time, there is a focal consumer who is performing the task on the platform. This focal consumer can see the online social signals generated by consumers who were on the platform performing tasks prior to him/her. It is also possible that several focal consumers are on the same platform at the same time performing the same task, but their co-existence on the platform would not interfere with each other, and each individual is still treated as a focal consumer. To distinguish these consumers, we use the ***focal consumer*** to refer to the consumer who is currently performing the task on the platform and is exposed to online social signals. We use ***prior consumers*** to refer to the consumers who were on the platform before the *focal consumer*, and whose information was used to generate online social signals for the *focal consumer*. Figure 2 presents the feedback loop of this process. To understand how online social signals are generated by using information from *prior consumers*, we need to delve deeper into how *prior consumers* go through the decision stages. Note that how consumers go through these decision stages are the same for both the *focal consumer* and *prior consumers*. Therefore, we will use "consumer" to discuss this general process.

Figure 2 Feedback Loop of Online Social Signals Generation Process



When consumers landed on a platform, they will start their decision-making journey, from the *information search* stage to the *post-consumption evaluation* stage (Dewey 1910, Engel et al. 1978). The details of these stages are discussed in the Introduction section.

Once consumers have finished all four decision stages (the *problem recognition* stage is not completed on the platform), they will leave the platform. However, the information consumers have generated on the platform will be stored at the backend. This information is a trail of data that consumers created while using the platform, which is also referred to as the digital footprints (Christensson 2014). This information is left online, either unintentionally or intentionally. The information that is left unintentionally can be a click on a link (Tucker and Zhang 2011), a listen of a song (Dewan and Ramaprasad 2012), a deal the consumer has purchased (Li and Wu 2018), etc. This information is left on the platforms without personally identifiable information. The information that is left intentionally on the platform is usually the information consumers submit on the platform. It can be an online review (Chevalier and Mayzlin 2006), a Thumbs Up for a video (Ameri et al. 2019), a comment under a post (Yang et al. 2019), etc. This information can either have personally identifiable information (e.g., names) or do not have personally identifiable information. Although the information from the consumers can be unintentionally or intentionally stored on the platform, the information that is unintentionally stored usually comes from consumer's *information search*, *alternative evaluation,* and *consumption* stage, and the information that is intentionally stored usually comes from the *post-consumption* evaluation stage.

The information or the traces each consumer left on the platform will then be processed by the platform at the backend. Once the platform has processed the information from *prior* *consumers*, it will then display these *online social signals* to *focal consumers* to aid their decision making on the platform.

While the above discussion focuses on how online social signals are created by prior consumers and processed by the platform, we have treated *prior consumers* equally. However, these *prior consumers* can have different relationships with the focal consumer. *Prior consumers* can be total strangers to the focal consumer, which can also be referred to as the crowd (Lee et al. 2015), or they can be the friends of focal consumers (Aral and Walker 2011). Prior consumers can also be experts in the field or public figures (Dellarocas et al. 2007).

Figure 3 Online Social Signals Generation Process

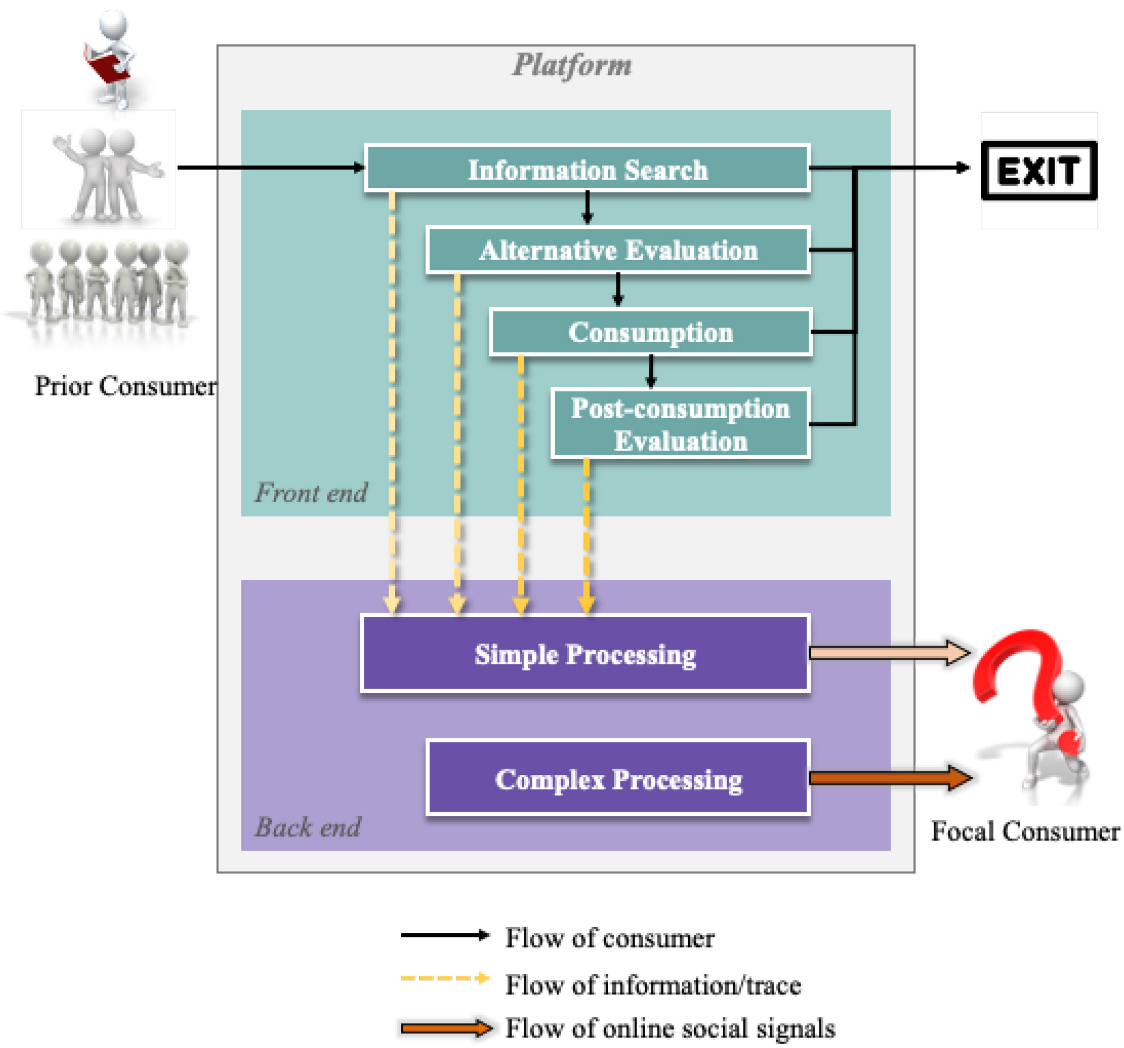


Figure 3 presents the whole process of how online social signals are generated on the platform. We use the frontend of the platform to denote the place where consumers interact with the platform and backend of the platform to refer to the place where information is being processed by the platform.

3. A Taxonomy of Online Social Signals

3.1. Dimensions of Online Social Signals

Based on the process of how online social signals are generated, we provide a taxonomy of online social signals that includes two dimensions.

The first dimension, ***information type***, refers to the information prior consumers created during a specific stage of their decision making. Information types include *search information*, *evaluation information*, *consumption information,* and *post-consumption information*. It is important to distinguish different types of online social signals as they can have a differential effect on consumer's decision making (Chen et al. 2011, Liu et al. 2017). *Search social signal* refers to the information that is generated from the prior consumer's information search stage. Prior research has used the number of clicks a website has received as a proxy for the search social signal (Tucker and Zhang 2011). *Evaluation social signal* refers to the information created from prior consumer's alternative evaluation stage when prior consumers are comparing the different alternatives they have. Amazon displays "customers who viewed this item also viewed" as a proxy for the evaluation social signal. *Consumption social signal* refers to the information generated by consumers' actions, which can include a consumption action or a purchase action (Dewan and Ramaprasad 2012, Qiu et al. 2018). Compared with the search social signal and evaluation social signal, consumption social signal contains more information as it suggests prior consumers have not only gone through the information search and alternative evaluation stages but have also made the consumption decision that includes both time and monetary effort. *Post-consumption social signal* is the information generated from prior consumer's opinions towards their consumption. It has been extensively studied in both the information systems literature and marketing literature (Rosario et al. 2016) in the form of online word-of-mouth and online review. The post-consumption social signal contains even more information compared with the consumption social signal as consumers technically have to first consume the product and then form their opinion towards the product. Therefore, post-consumption social signals reveal consumer's opinions and recommendations about a product with reasoning (Chen et al. 2011). While we have discussed four information types, prior studies have mostly focused on consumption social signal and post-consumption social signal. Limited studies have examined the search social signal and evaluation social signal. To simplify our conceptual framework, we use the *pre-consumption social signal* to refer to the type of social signals from search, evaluation, and consumption.

The second dimension, ***information source***, is also critical because the same social signals coming from different sources can have a differential impact (Dewan et al. 2017, Gu et al. 2012). We epitomize this dimension through three categories: *crowd*, *friends,* and *experts*. When the social signal comes from the *crowd*, it also means the information comes from prior consumers with whom the focal consumer does not personally have a connection. Most of the prior studies in Table 1 have focused on studying the social signals generated by the crowd. Social signals from *friends* are different from the social signals from the crowd as it can contain more private information and can be interpreted better by the focal consumer based on their knowledge about their friends (Lee et al. 2015). Note that in the internet world, this online "friend" does not necessarily have to be a friend in the offline world. Social signals from *experts* are broadly defined, which can be from professional experts or third-party sources. Expert social signals can contain more information compared with crowd social signals because it is more specialized, which offers greater depth and insights (Gu et al. 2012).

3.2. Online Social Signal Subtype

The combination of the two dimensions yields a 2 x 3 taxonomy with five online social signal types [1]. Table 2 provides the definition and practical examples of these online social signals.

3.2.1. Crowd's pre-consumption signal

Crowd's pre-consumption signal refers to the actions or decisions taken by prior consumers on a platform. When consumers have imperfect information about the product, they tend to rely on prior consumer's actions or decisions (Banerjee 1992, Bikhchandani et al. 1992, 1998). Consumers learn from these observed signals and update their own beliefs about the product. This process is often called observational learning (Cai et al. 2009). More specifically, information cascade refers to the phenomena where consumers follow prior consumer's decisions and ignore their private information, and herding behavior refers to the phenomena where consumers are making identical changes without necessarily ignoring their private information (Çelen and Kariv 2004). Through the observational learning process, consumers draw quality inferences from observing the choices of other consumers to reduce the uncertainty they are facing. Therefore, the crowd's pre-consumption signal can effectively influence the consumer's decisions. Crowd's pre-consumption signal can be particularly prominent in the decision making because online platforms can easily display these signals to the focal consumer

---

[1] While expert pre-consumption social signal exists theoretically, prior studies have not examined this type and we did not find practical examples of this social signal. Therefore, we do not discuss this type in our paper.

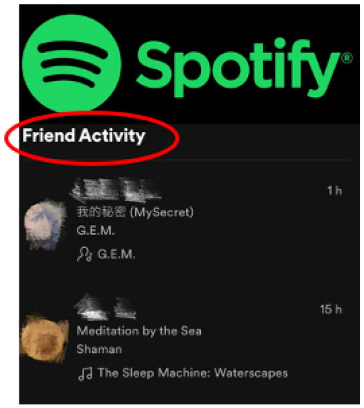
Table 1 Prior Literature on Online Social Signals and Consumer Decision

	Online Social Signals						Consumer Decision		
Information Source	Crowd		Friend		Expert		Search	Choice	Post-
Information Type	Pre-	Post-	Pre-	Post-	Pre-	Post-			
Papers that only examine one social signal									
Tucker and Zhang (2011)	✓						✓		
Dewan and Ramaprasad (2012)	✓						✓		
Oestreicher-Singer and Sundararajan (2012)	✓							✓	
Dewan and Ramaprasad (2014)	✓							✓	
Aggarwal and Singh (2013)	✓						✓	✓	
Chevalier and Mayzlin (2006)		✓						✓	
Forman, Ghose, Wiesenfeld (2008)		✓						✓	
Li and Hitt (2008)		✓						✓	
Aggarwal, Gopal, Gupta, Singh (2012)		✓						✓	
Lu, Ba, Huang, Feng (2013)		✓						✓	
Godes and Mayzlin (2004)		✓							✓
Li (2018)		✓							✓
Adomavicius, Bockstedt, Curley and Zhang (2013)		✓							✓

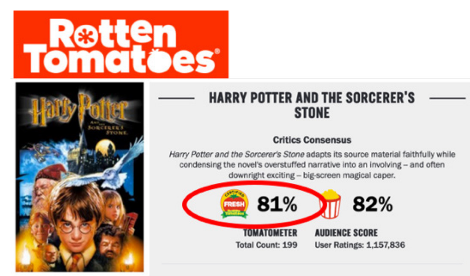
	Online Social Signals						Consumer Decision		
Information Source	Crowd		Friend		Expert		Search	Choice	Post-
Information Type	Pre-	Post-	Pre-	Post-	Pre-	Post-			
Yang, Ren, Adomavicius (2018)		✓							✓
Huang, Sun, Chen, Golden (2019)		✓					✓	✓	
Bapna and Umyarov (2015)			✓					✓	
Qiu, Shi Whinston (2018)			✓					✓	
Wang, Zhang, Hann (2018)				✓					✓
Luo, Gu, Zhang, Phang (2017)						✓			✓
Papers that examine more than one social signals									
Duan, Gu, Whinston (2009)	✓	✓						✓	
Chen, Wang, Xie (2011)	✓	✓						✓	
Lee, Hosanagar, Tan (2015)		✓		✓					✓
Dewan, Ho, Ramaprasad (2017)		✓		✓			✓		
Gu, Park, Konana (2012)		✓				✓		✓	
Huang, Boh, Goh		✓				✓		✓	
Li and Wu (2018)	✓			✓				✓	

Table 2 Subtypes of Online Social Signals

<table>
<tr><td></td><td></td><th colspan="3">Information Source</th></tr>
<tr><td></td><td></td><td>Crowd</td><td>Friend</td><td>Expert</td></tr>
<tr><td rowspan="2">Information Type</td><td>Pre-consumption</td><td>Actions or decisions from the crowd.

</td><td>Actions or decisions from the focal consumer’s friends.

</td><td>N/A</td></tr>
<tr><td>Post-consumption</td><td>Opinions from the crowd.

</td><td>Opinions from the focal consumer’s friends.

</td><td>Opinions from the expert.

</td></tr>
</table>

(Duan et al. 2009). Given the increasing amount of information online (Brynjolfsson and Smith 2000), online consumers may also find it more efficient to rely on these crowd's pre-consumption signals as it can save their time and effort. Examples of crowd's pre-consumption signals include sales rank and number of sales (Li and Wu 2018), blog mentions (Dewan and Ramaprasad 2012), the total number of downloads (Duan et al. 2009), and Amazon's "what do consumers ultimately buy after viewing this item" (Chen et al. 2011) and "customers who bought this item also bought" (Oestreicher-Singer and Sundararajan 2012).

3.2.2. Crowd' post-consumption signal

Crowd's post-consumption signal represents the act of consumers providing opinions about goods, services, brands, or companies to other consumers at the online platforms (Rosario et al. 2016). It is also known as online review, user-generated content, or electronic word-of-mouth. This post-consumption signal encompasses the crowd's knowledge towards a product in terms of their own experience, recommendations, or complaints (Kannan and Li 2017). The post-consumption signal is created after consumers have experienced it, therefore implicitly assuming consumers have already gone through the consumption stage. Similar to traditional (offline) word-of-mouth that is influential for consumer's decision making (Katz and Lazarsfeld 1966), the online crowd's post-consumption signal is also very effective (Chevalier and Mayzlin 2006, Godes and Mayzlin 2004). Crowd's post-consumption signal not only convey the existence of a product (awareness effect) but also shape consumers' attitude and evaluation towards the product (persuasive effect) that ultimately will change consumer's decision making (Duan et al. 2008a). Examples of crowd's post-consumption signals include volume, valence and variance of online reviews and online posts (Chevalier and Mayzlin 2006, Godes and Mayzlin 2004, Li 2018, Sun 2012, Yang et al. 2019), reviewer identity (Forman et al. 2008), number of "Likes" of the product or posts (Dewan et al. 2017, Li and Wu 2018), and the actual text and emotions embedded in the signal (Büschken and Allenby 2016, Humphreys and Wang 2018).

3.2.3. Friend's pre-consumption signal

Friend's pre-consumption signals are very similar to the crowd's pre-consumption signals in that consumers are likely to follow their friend's decisions because of the uncertainty involved in making choices, which is the observational learning mechanism. However, given the sources of these two pre-consumption signals are different (crowd vs. friend), two other mechanisms also explain why consumers are influenced by their friend's decisions (Cai et al. 2009). The first mechanism is the conformity effect or peer pressure (Asch 1951), in which consumers follow their friends' decisions either because they want to form an accurate decision or because they want to obtain social approval from others (Cialdini and Goldstein 2004). Consumers are likely to conform to their friends by changing their behavior to fit in with others. The second mechanism is the homophily effect (McPherson et al. 2001). People tend to be friends with others who share similar sociodemographic, behavioral, and intrapersonal characteristics. Even though focal consumers may not be influenced by their friends, but they will make similar decisions because of their intrinsic similarity (Bapna and Umyarov 2015). Regardless of the mechanisms that drive the focal consumer's decision, platforms would observe consumers are making similar decisions provided the availability of their friend's pre-consumption signal (Qiu et al. 2018). Examples of friends' pre-consumption signal can include friend's "check-in" action in the restaurant app (Qiu et al. 2018), and friend's subscription to online service (Bapna and Umyarov 2015).

3.2.4. Friend's post-consumption signal

Friend's post-consumption signal encompasses the opinion and experience information from prior consumers who are friends with the focal consumer. Friend's post-consumption signal can not only increase product awareness but also can reduce the quality uncertainty related to the product because it comes from online social ties that have similar tastes or know about the consumer's idiosyncratic preferences (McPherson et al. 2001). Consumers can also interpret the signals from their friends better based on their knowledge about their friends (Li and Wu 2018). Friend's post-consumption signals can carry more private information because of the communication tools. Friends can leverage the social tools and functions on the platform to send private messages or chat with each other (Ameri et al. 2019, Lee et al. 2015). Even after accounting for the homophily among friends, the friend's post-consumption signal still significantly influence focal consumers' behavior, and it is stronger for the focal consumer who has smaller networks (Wang et al. 2018, Zhang et al. 2015). Examples of friend's post-consumption signal include average prior rating by friends for a given product (Ameri et al. 2019, Lee et al. 2015, Wang et al. 2018), friend's favoring behavior (Dewan et al. 2017), and post's cumulative likes and tweets (Li and Wu 2018).

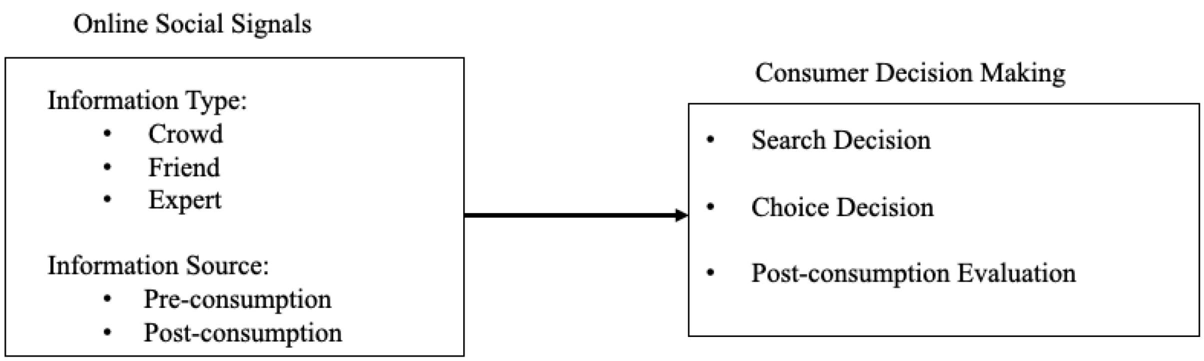
3.2.5. Expert's post-consumption signal

Expert's post-consumption signal refers to the opinion experts provide about goods, services, brands, or companies to other consumers at the online platforms. These experts are usually employed by the platform to provide their subjective views or comments on the products (Eliashberg and Shugan 1997). Thus, the post-consumption signals from experts tend to contain more product information and industry insights compared with the signals from the crowd (Gu et al. 2012, Luo et al. 2017). These experts are usually specialized in that product category so that their expertise can influence the persuasiveness of the opinion they give out (Huang et al. 2017). The opinion from these experts or critics plays a significant role in consumer's decision making from choosing which stock to purchase (Goh and Ederington 1993) to what movie to watch (Basuroy et al. 2003, Eliashberg and Shugan 1997, Zhang and Dellarocas 2006). Examples of expert's post-consumption signals mostly include review and ratings given by the experts.

4. Conceptual Framework of Online Social Signals on Consumer Behavior

Consumers go through different decision stages in the decision-making process because they want to reduce the perceived risk and make a better decision (Cox 1967). Information processing theory also suggests that consumer decision making can take multiple stages because of the limited capacity to process information (Bettman 1979, Payne 1982). Online platforms have provided consumers with much more information compared with the offline environment. In a way, it can help consumers reduces uncertainty and perceived risks for their decision-making process. However, consumers usually make a trade-off between the accuracy of the decision and the effort required to make the decision (Payne 1982). Even though the abundant amount of information online can potentially increase consumer's decision accuracy, it has also drastically increased consumer's effort as they need to process more information. Therefore, consumers are more likely to follow other people's decisions to reduce the decision effort and increase decision accuracy when they have imperfect private information. Online social signals, a digitized way to present prior consumer's decision making, are important factors in affecting consumer's decision-making. While online social signals are created by prior consumers, their influences are on the *focal consumer*'s decision making. Similar to prior consumers, focal consumer's decision making on the platform also consists of four stages from *information search* stage *to post-consumption* stage (*problem recognition* stage cannot be observed online). It is important to understand how online social signals would influence these decision stages. Do online social signals play the same role for each decision stage or they have differential impact on different stages? To further understand this question, we propose a conceptual framework (Figure 4) that examines the influence of online social signals on consumer decision making by exploring focal consumer's *search* decision, *consumption* (i.e. choice) decision and *post-consumption* evaluation[2].

Figure 4 Conceptual Framework



4.1. Search Decision

In the online environment, consumers might experience higher uncertainty compared to offline because they cannot physically experience the products (Dimoka et al. 2012). Therefore, they have the motivation to search for information (proxy by consumer's " click " behavior) to reduce the perceived risk and to make better decisions. When consumers are at the *information search* stage, they only have a broad and fuzzy idea of they want to purchase or consume (Lambrecht and Tucker 2013). They are not consciously aware of their category needs or brand preferences. Only after clicking on a product or link, consumers can obtain additional information about the product and better construct their preferences. Given consumers are still

[2] The *alternative evaluation* decision is hard to observe online and studies have shown that consumers sometime does not always perform this decision (Olshavsky and Granbois 1979).

unaware of their true preference before the *search* decisions and given the uncertainty and imperfect information associated with making decisions online, consumers are highly susceptible to other influences and are very receptive to advice (Bleier and Eisenbeiss 2015). Therefore, online social signals will have a strong influence on consumers' *search* decisions when consumers only have an abstract and high-level of preferences (Dewan et al. 2017, Dewan and Ramaprasad 2012, Tucker and Zhang 2011).

4.2. Consumption Decision

After focal consumers have searched and evaluated the information, they make the *consumption* decisions. Through the previous decision stages, consumers have already obtained some information about the products. Therefore, their uncertainty or perceived risk associated with the *consumption* decision has decreased compared with when they are at the *information search* stage. Some consumers may think the information they have obtained is already enough for them to make the decisions, implying they are less dependent on online social signals for *consumption* decisions because of the increased preference stability (Bleier and Eisenbeiss 2015). Other consumers may still face uncertainties as the information or knowledge they acquire from the *information search* stage is not equivalent to the full experience they get after consumption. The Internet interface has set up a barrier in experiencing products' experience qualities that prevent these consumers to fully get rid of the uncertainties (Dimoka et al. 2012, Nelson 1970, 1974). Therefore, these "uncertain consumers" will still rely on online social signals in the *consumption* stage. However, the *consumption* decision can involve other factors that go beyond the information acquired from online social signals (Huang et al. 2019), such as shipping cost (Lewis et al. 2006). Even if the "uncertain consumers" use online social signals for the *consumption* decision, online social signals are not the only factor that influences their decisions. Thus, overall, online social signals may either have no effect or a weak effect on the *consumption* decision after considering the *search* decision (Aggarwal and Singh 2013, Huang et al. 2019). Without considering the *search* decision, we may attribute the effect of online social signals on *search* decision entirely to the effect of online social signals on *consumption* decision and observe a strong impact on sales or individual consumption decisions (Aggarwal et al. 2012, Chevalier and Mayzlin 2006, Duan et al. 2009, Li and Wu 2018).

4.3. Post-consumption Evaluation

Consumers go through the post-consumption stage after they have made the *consumption* decision, meaning they have already experienced the product and have individual opinions towards the product. Although consumers already have a self-perceived quality of the product, they might still be influenced by the online social signals (specifically, the post-consumption signals). Two mechanisms can explain this. The first one is the conformity effect (peer pressure). While consumers face little uncertainties in the *post-consumption* stage compared with previous stages, they might still have difficulty in evaluating some features of the products or services even after consumption (Dulleck and Kerschbamer 2006, Wang et al. 2018). Therefore, the focal consumer may decide to conform with other people's opinions by changing or ignoring their own opinion after observing prior consumer's post-consumption signals (Banerjee 1992, Bikhchandani et al. 1992, Cialdini and Goldstein 2004). The focal consumer might think it is optimal and rational to follow prior consumers' *post-consumption* decisions with the fear of losing reputation (Lee et al. 2015). The second mechanism is the anchoring effect. The post-consumption signals from prior consumers may serve as an anchor for the focal consumer, which results in the focal consumer's *post-consumption decision* becomes bias towards this anchor that they first observe (Adomavicius et al. 2013, Tversky and Kahneman 1974). Therefore, online social signals can still influence consumer's *post-purchase* evaluation even after consumption (Lee et al. 2015, Wang et al. 2018, Yang et al. 2019).

5. Discussions

We consider the main effect of online social signals on consumer decision-making in the conceptual framework (Figure 4). However, the effect of online social signals can be heterogeneous. In this section, we explore two sources of heterogeneity: product characteristics and consumer characteristics.

We examine three characteristics of the product. First, we account for different types of products: search products, experience products, and credence products. Search products refer to the products whose attribute quality can be easily obtained prior to consumption.[3] Experience products refer to the products whose attribute quality can only be acquired after consumption (Nelson 1970, 1974). Credence products are products whose attribute quality is difficult or impossible to evaluate even after consumption (Darby and Karni 1973, Dulleck and Kerschbamer 2006). Among these three types of products, credence products have the largest product uncertainty because it is most difficult for consumers to evaluate credence products and to predict how it can perform in the future, and search products have the smallest uncertainty (Dimoka et al. 2012). For consumers' *search* decisions, the effect of online social signals will not be different across product types because online social signals all serve the purpose of reducing consumer's quality uncertainty at this stage. However, it will have a differential impact on consumer's *consumption* decisions depending on the product type. It would be more effective for experience goods and credence goods compared with search goods because online social signals

[3] Products are made up of different attributes. Therefore, search goods can have search attributes, experience attributes and credence attribute. However, the dominant attributes of search goods are search attributes (Nelson 1974). Similar logic applies to experience goods and credence goods.

can serve as proxies for those qualities that cannot be assessed after the *search* stage (Li and Wu 2018, Wang et al. 2018). However, if experience products can be "sampled" at the *information search* stage, then online social signals might not be as influential for the *consumption* decision as they are without sampling option (Jin et al. 2019). Consumers can "virtually experience" the experience product prior to their consumption, thus turning the product into search products (Klein 1998). In this case, the influence of online social signals on *consumption* decisions can be similar for both search products and experience products, which is either very weak or no effect.

Second, products can vary based on whether they are mainstream products or niche products, the long-tail phenomenon brought by the proliferation of online platforms (Anderson 2006). On the one hand, a mainstream product means it serves the mainstream taste and is more likely to be preferred compared with niche product if the quality is the same (Tucker and Zhang 2011). Consumers might be more aware of mainstream products compared with niche products. Therefore, the need to use online social signals to reduce risk is low for mainstream products. On the other hand, niche products usually contain a higher level of quality uncertainty because fewer consumers have been exposed to the product (Dewan and Ramaprasad 2012). Online social signals for niche products will be more salient because of the scarcity of available information about the niche product (Zhu and Zhang 2010). Given the differences between mainstream products and niche products, online social signals will have a stronger influence for niche products than the mainstream product (Dewan et al. 2017, Dewan and Ramaprasad 2012, Tucker and Zhang 2011, Zhu and Zhang 2010).

Third, products can present network externalities, which is the effect of an additional user can create on the total value of the product (Asvanund et al. 2004). Take mobile apps as an example. A messenger app has a network externality because the more people use the app, the more convenient it is. In contrast, a weather app does not have a network externality because additional consumers using it does not provide more benefits for the current users. Social signals can be seen as a proxy for how many consumers are using the product, which thus influences the consumer's *consumption* decision of network externalities products.

Different consumer characteristics can also result in a heterogeneous impact of online social signals. The first one is product expertise. When consumers first encounter a new product, they are more likely to be constructing their preferences on the product, and their preferences are unstable (Hoeffler and Ariely 1999). Therefore, when consumers do not have much expertise in the product, they are more likely to rely on online social signals for their decision-making because of the product uncertainties. However, as consumers get more familiar with the product and gain more experience, their preferences become more stable. Hence, online social signals will be less effective for experienced consumers as they have already constructed their own opinion about the product and have fewer product uncertainties compared with consumers with low product expertise. The second characteristic is platform expertise. Prior research has shown that consumers with higher internet experience are less likely to be influenced by online social signals because they are more familiar with the online channels and have less perception of uncertainty and complexity (Zhu and Zhang 2010). While online shopping has become much more popular compared with ten years ago, and more consumers are familiar with the Internet, Internet experience is not very related to nowadays consumer's decisions. However, on a similar level, platform expertise can be an important moderator. Consumers with a high level of platform expertise will feel less uncertain and complex during their decision-making process and will be less influenced by online social signals. Consumers with low-level platform expertise will have less confidence and evoke the perception of uncertainty when making their decisions. Therefore, they are more likely to be influenced by online social signals during their decision-making process.

6. Conclusion

In this paper, we present a taxonomy on online social signals and a conceptual framework on the influence of online social signals on consumer decision making. By analyzing how online social signals are generated fundamentally, we are able to propose two dimensions of social signals and discuss each type of social signals in detail. Our conceptual framework further discusses the differential impact of social signals on different decision stages.

This study contributes to the literature in several ways. First, we provide a taxonomy to systematically determine different types of online social signals with two dimensions: *information type*, and *information source*. For online social signals determined by these two components, we analyze how each of these online social signals is different and what are the underlying mechanisms that made these social signals influential to consumers. Second, we provide a conceptual framework by theorizing how online social signals influence focal consumer's decisions differently: search, consumption, and post-consumption. This conceptual framework, along with the taxonomy, provides an adequate outline to account for the phenomena relating to online social signals. They can help us effectively synthesize the findings from prior research and provide a theoretical foundation in comparing different online social signals and different decision stages. In addition, this framework can also help us in identifying the research gaps. This study also offers significant insights for the practitioners. Online platforms can use our frameworks as a roadmap for designing different types of online social signals based on their needs and the product and consumers they are encountering.

Our study is not without limitations. First, our framework does not account for the online social signals from different platforms. It would be very interesting to take into account this factor and examine how cross-platform signals can have different strengths. Second, our framework does not differentiate the different types of online platforms, such as e-commerce platforms, review platforms, and social media platforms. Different types of platforms can have fundamental differences, including the types of products on the platform, the types of consumers on the platform, and the type of need when consumers visit the platform. Exploring online social signals' roles in different types of platforms is an interesting extension of this study. Third, our current framework cannot classify one type of consumer decision, which is the decision to give an opinion on a post-consumption evaluation. It can take the form of saying a review is helpful or a comment that replies to a previous comment. Since it is hard to identify whether consumers give out this opinion before their *consumption* decision or after their *consumption* decision, we do not include this decision in our framework. Future research can theorize this decision base on the types of platforms (e.g., consumers generally "like" a review before their *consumption* decision when they are on e-commerce platforms and review platforms, but they may "comment" to a comment on social media platform after their consumption) to enrich the current framework.

7. References

Adomavicius G, Bockstedt JC, Curley SP, Zhang J (2013) Do Recommender Systems Manipulate Consumer Preferences? A Study of Anchoring Effects. *Inf. Syst. Res.* 24(4):956–975.

Aggarwal R, Gopal R, Gupta A, Singh H (2012) Putting money where the mouths are: The relation between venture financing and electronic word-of-mouth. *Inf. Syst. Res.* 23(3-part-2):976–992.

Aggarwal R, Singh H (2013) Differential influence of blogs across different stages of decision making: The case of venture capitalists. *MIS Q.* 37(4):1093–1112.

Ameri M, Honka E, Xie Y (2019) Word of mouth, observed adoptions, and anime-watching decisions: The role of the personal vs. the community network. *Mark. Sci.* 38(4):567–583.

Anderson C (2006) *The long tail: Why the future of business is selling less of more* (Hachette Books).

Aral S, Walker D (2011) Creating social contagion through viral product design: A randomized trial of peer influence in networks. *Manage. Sci.* 57(9):1623–1639.

Asch SE (1951) *Effects of group pressure on the modification and distortion of judgments*

Asvanund A, Clay K, Krishnan R, Smith MD (2004) An empirical analysis of network externalities in peer-to-peer music-sharing networks. *Inf. Syst. Res.* 15(2):155–174. http://pubsonline.informs.org.https//doi.org/10.1287/isre.1040.0020http://www.informs.org.

Banerjee A V. (1992) A Simple Model of Herd Behavior. *Q. J. Econ.* 107(3):797–817.

Bapna R, Umyarov A (2015) Do your online friends make you pay? A randomized field experiment on peer influence in online social networks. *Manage. Sci.* 61(8):1902–1920.

Basuroy S, Chatterjee S, Abraham Ravid S (2003) How Critical are Critical Reviews? The Box

Office Effects of Film Critics, Star Power, and Budgets. *J. Mark.* 67(4):103–117.

Bettman JR (1979) *An Information Processing Theory of Consumer Choice*

Bikhchandani S, Hirshleifer D, Welch I (1992) A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *J. Polit. Econ.* 100(5):992–1026.

Bikhchandani S, Hirshleifer D, Welch I (1998) Learning from the behavior of others: Conformity, fads, and informational cascades. *J. Econ. Perspect.* 12(3):151–170.

Bleier A, Eisenbeiss M (2015) Personalized online advertising effectiveness: The interplay of what, when, and where. *Mark. Sci.* 34(5):669–688.

Brynjolfsson E, Smith MD (2000) Frictionless commerce? A comparison of Internet and conventional retailers. *Manage. Sci.* 46(4):563–585.

Büschken J, Allenby GM (2016) Sentence-based text analysis for customer reviews. *Mark. Sci.* 35(6):953–975. (November 1).

Cai H, Chen Y, Fang H (2009) Observational learning: Evidence from a randomized natural field experiment. *Am. Econ. Rev.* 99(3):864–882.

Çelen B, Kariv S (2004) Distinguishing informational cascades from herd behavior in the laboratory. *Am. Econ. Rev.* 94(3):484–498.

Chen Y, Wang Q, Xie J (2011) Online Social Interactions: A Natural Experiment on Word of Mouth Versus Observational Learning. *J. Mark. Res.* 48(2):238–254.

Chevalier JA, Mayzlin D (2006) The Effect of Word of Mouth on Sales: Online Book Reviews. *J. Mark. Res.* 43(3):345–354.

Chintagunta PK, Gopinath S, Venkataraman S (2010) The Effects of Online User Reviews on Movie Box Office Performance: Accounting for Sequential Rollout and Aggregation Across Local Markets. *Mark. Sci.* 29(5):944–957.

Christensson P (2014) Digital Footprint Definition. *TechTerms*.

Cialdini RB, Goldstein NJ (2004) Social Influence: Compliance and Conformity. *Annu. Rev. Psychol.* 55(1):591–621.

Cox DF (1967) *Risk-Taking and Information Handling in Consumer Behavior* (Harvard University Press).

Darby MR, Karni E (1973) Free Competition and the Optimal Amount of Fraud. *J. Law Econ.* 16(1):67–88.

Dellarocas C, Zhang X, Awad NF (2007) Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *J. Interact. Mark.* 21(4):23–45.

Dewan S, Ho YJ, Ramaprasad J (2017) Popularity or proximity: Characterizing the nature of social influence in an online music community. *Inf. Syst. Res.* 28(1):117–136.

Dewan S, Ramaprasad J (2012) Music blogging, online sampling, and the long tail. *Inf. Syst. Res.* 23(3 PART 2):1056–1067.

Dewey J (1910) *How we think* (D.C. Heath & Co., Boston, Mass. SE - vi, 224 pages ; 20 cm, Mass. SE - vi, 224 pages ; 20 cm).

Dimoka A, Hong Y, Pavlou PA (2012) On product uncertainty in online markets: Theory and evidence. *MIS Q.* 36(2):395–426.

Duan W, Gu B, Whinston AB (2008a) Do online reviews matter? - An empirical investigation of panel data. *Decis. Support Syst.* 45(4):1007–1016.

Duan W, Gu B, Whinston AB (2008b) The dynamics of online word-of-mouth and product sales-An empirical investigation of the movie industry. *J. Retail.* 84(2):233–242.

Duan W, Gu B, Whinston AB (2009) Informational cascades and software adoption on the Internet: An empirical investigation. *MIS Q.* 33(1):23–48.

Dulleck U, Kerschbamer R (2006) On doctors, mechanics, and computer specialists: The economics of credence goods. *J. Econ. Lit.* 44(1):5–42. https://www.aeaweb.org/articles?id=10.1257/002205106776162717.

Eliashberg J, Shugan SM (1997) Film Critics: Influencers or Predictors? *J. Mark.* 61(2):68–78.

Engel JF, Kollat DT, Blackwell RD (1978) *Consumer Behavior* (The Dryden Press).

Forman C, Ghose A, Wiesenfeld B (2008) Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Inf. Syst. Res.* 19(3):291–313. (September) http://pubsonline.informs.org/doi/10.1287/isre.1080.0193.

Girardin F, Blat J, Calabrese F, Dal Fiore F, Ratti C (2008) Digital footprinting: Uncovering tourists with user-generated content. *IEEE Pervasive Comput.* 7(4):36–44.

Godes D, Mayzlin D (2004) Using online conversations to study word-of-mouth communication. *Mark. Sci.* 23(4):545–560.

Goh JC, Ederington LH (1993) Is a Bond Rating Downgrade Bad News, Good News, or No News for Stockholders? *J. Finance* 48(5):2001–2008.

Gu B, Park J, Konana P (2012) The impact of external word-of-mouth sources on retailer sales of high-involvement products. *Inf. Syst. Res.* 23(1):182–196.

Hoeffler S, Ariely D (1999) Constructing stable preferences: A look into dimensions of experience and their impact on preference stability. *J. Consum. Psychol.* 8(2):113–139.

Huang J, Boh WF, Goh KH (2017) A Temporal Study of the Effects of Online Opinions: Information Sources Matter. *J. Manag. Inf. Syst.* 34(4):1169–1202.

Huang N, Sun T, Chen P, Golden JM (2019) Word-of-Mouth System Implementation and Customer Conversion: A Randomized Field Experiment. *Inf. Syst. Res.* 30(3):805–818.

Humphreys A, Wang RJH (2018) Automated text analysis for consumer research. *J. Consum.*

Res. 44(6):1274–1306.

Jin Q, Animesh A, Pinsonneault A (2019) When popularity meets position. *40th Int. Conf. Inf. Syst. ICIS 2019*.

Kannan PK, Li H "Alice" (2017) Digital marketing: A framework, review and research agenda. *Int. J. Res. Mark.* 34(1).

Katz E, Lazarsfeld PF (1966) *Personal influence: The part played by people in the flow of mass communications*

Klein LR (1998) Evaluating the Potential of Interactive Media through a New Lens: Search versus Experience Goods. *J. Bus. Res.* 41(3):195–203.

Lambrecht A, Tucker C (2013) When does retargeting work? Information specificity in online advertising. *J. Mark. Res.* 50(5):561–576.

Lee YJ, Hosanagar K, Tan Y (2015) Do I Follow My Friends or the Crowd? Information Cascades in Online Movie Ratings. *Manage. Sci.* 61(9):2241–2258.

Lewis M, Singh V, Fay S (2006) An empirical study of the impact of nonlinear shipping and handling fees on purchase incidence and expenditure decisions. *Mark. Sci.* 25(1):51–64.

Li X (2018) Impact of average rating on social media endorsement: The moderating role of rating dispersion and discount threshold. *Inf. Syst. Res.* 29(3):739–754.

Li X, Wu L (2018) Herding and Social Media Word-of-Mouth: Evidence From Groupon. *MIS Q.* 42(4):1331.

Liu Y, Feng J, Liao X (2017) When online reviews meet sales volume information: Is more or accurate information always better? *Inf. Syst. Res.* 28(4):723–743.

Luo X, Gu B, Zhang J, Phang CW (2017) Expert blogs and consumer perceptions of competing brands. *MIS Q.* 41(2):371–395.

McPherson M, Smith-Lovin L, Cook JM (2001) Birds of a Feather: Homophily in Social Networks. *Annu. Rev. Sociol.* 27(1):415–444.

Merriam-Webster Consume | Definition of Consume by Merriam-Webster. Retrieved (June 12, 2020), https://www.merriam-webster.com/dictionary/consume.

Nelson P (1970) Information and Consumer Behavior. *J. Polit. Econ.* 78(2):311–329.

Nelson P (1974) Advertising as Information. *J. Polit. Econ.* 82(4):729–754.

NPR (2018) NPR/Marist Poll: Amazon is a colossus in a nation of shoppers : NPR. *NPR* https://www.npr.org/about-npr/617470695/npr-marist-poll-amazon-is-a-colossus-in-a-nation-of-shoppers.

Oestreicher-Singer G, Sundararajan A (2012) Recommendation networks and the long tail of electronic commerce. *MIS Q.* 36(1).

Olshavsky RW, Granbois DH (1979) Consumer Decision Making-Fact or Fiction? *J. Consum. Res.* 6(2).

Payne JW (1982) Contingent decision behavior. *Psychol. Bull.* 92(2):382–402.

Qiu L, (Michael) Shi Z, Whinston AB (2018) Learning from your friends' check-ins: An empirical study of location-based social networks. *Inf. Syst. Res.* 29(4):1044–1061.

Rosario AB, Sotgiu F, De Valck K, Bijmolt THA (2016) The effect of electronic word of mouth on sales: A meta-analytic review of platform, product, and metric factors. *J. Mark. Res.* 53(3):297–318.

Rouse M (2014) Digital Footprint.

Spence M (1978) Job market signaling. *Uncertain. Econ.* (Elsevier), 281–306.

Sun M (2012) How Does the Variance of Product Ratings Matter? *Manage. Sci.* 58(4):696–707.

Tucker C, Zhang J (2011) How Does Popularity Information Affect Choices? A Field

Experiment. *Manage. Sci.* 57(5):828–842.

Tversky A, Kahneman D (1974) Judgment under uncertainty: heuristics and biases. *Science (80-.).* 185:1124–1131.

Wang C (Alex), Zhang X (Michael), Hann IH (2018) Socially nudged: A quasi-experimental study of friends' social influence in online product ratings. *Inf. Syst. Res.* 29(3):641–655.

Yang M, Ren YC, Adomavicius G (2019) Understanding User-Generated Content and Customer Engagement on Facebook Business Pages. *Inf. Syst. Res.* 30(3):839–855.

Zhang J, Liu P (2012) Rational Herding in Microloan Markets. *Manage. Sci.* 58(5):892–912.

Zhang J, Liu Y, Chen Y (2015) Social learning in networks of friends versus strangers. *Mark. Sci.* 34(4):573–589.

Zhang X (Michael), Dellarocas C (2006) The Lord Of The Ratings: Is A Movie's Fate Influenced by Reviews? *ICIS 2006 Proc.*

Zhu F, Zhang X (Michael) (2010) Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics. *J. Mark.* 74(2):133–148.

Chapter 3. Does Popularity Really Matter—Disentangling Popularity and Position Effect Using an Experimental Approach

1. Introduction

The Internet has made it easier for consumers to overcome geographical and temporal barriers to purchase or access products anytime, anywhere. It has also led to the availability of a large number of alternatives. For example, Amazon sells more than 12 million products (Dayton 2020), Spotify has over 50 million songs on their platform (Spotify 2020), and YouTube has more than 720,000 hours of video are uploaded to the platform every day (Mohsin 2020). While platforms are offering more alternatives for consumers to choose from, they are also providing more information on each alternative as the cost to distribute the information is low. For example, BestBuy displays its electronics with all detailed specifications, and Netflix presents each movie with its description and detail casting information. However, the increased amount of available information to the consumers can potentially lead to information overload issues and result in poorer decisions (Jacoby et al. 1974, Jeffries 2015). To make it easier for the consumers to manage this possible information overload issue, platforms have started to display quality signals along with the regular descriptive information to help consumers make better decisions. Among these quality signals, **popularity information** is widely used, which refers to *the aggregated decisions or activities from prior consumers or prior users* (Dewan et al. 2017, Tucker and Zhang 2011, Zhu and Zhang 2010).

Platforms have different ways to display this popularity information. Figure 5 provides an example of how the popularity information is shown on a music platform. Popularity, which is proxy by the daily plays, is shown in descending order, and the most popular song (i.e., the song that has the largest daily plays) is on the top position. The more popular the song is, the higher the position it has on the platform. However, not all the platforms sort products by their popularity. Figure 6 provides another example from a music platform where popularity information is not sorted in the descending order (i.e., unsorted)[4]. Although the song at the top position in this example is the most popular one, the song on the third position is not accordingly the third most popular song. While platforms have used different ways to display the popularity information to help consumers' decision processes, research has also shown that the popularity information does have an influence on consumers' decision processes.

Figure 5 Example of Platform where Popularity is in Descending Order

TITLE		ARTIST	DAILY PLAYS
Blinding Lights		The Weeknd	4,516,393
ROCKSTAR (feat. Roddy Ricch)	EXPLICIT	DaBaby, Roddy Ricch	4,050,048
Roses - Imanbek Remix	EXPLICIT	SAINt JHN, Imanbek	3,793,500
Rain On Me (with Ariana Grande)		Lady Gaga, Ariana Grande	3,270,819
Toosie Slide	EXPLICIT	Drake	3,171,789
death bed (coffee for your head) (feat. beabadoobee)		Powfu, beabadoobee	3,127,940

Figure 6 Example of Platform where Popularity is Unsorted

NAME		ARTIST	ALBUM	TIME	POPULARITY
Tommy Lee (feat. Post Malone)	E	Tyla Yaweh	Tommy Lee (feat. Post Malone) - Single	3:44	
Both Still Young		Dzeko & Keith Urban	Both Still Young - Single	2:53	
Learn To Fly		Surfaces & Elton John	Learn To Fly - Single	3:29	
Summer Time		James Barker Band	Summer Time - Single	3:06	
Don't Rush (feat. DaBaby)	E	Young T & Bugsey	Don't Rush (feat. DaBaby) - Single	3:21	
Dead Inside		Lo Lo	Dead Inside - Single	2:47	

With limited information available, people tend to follow other people's actions or

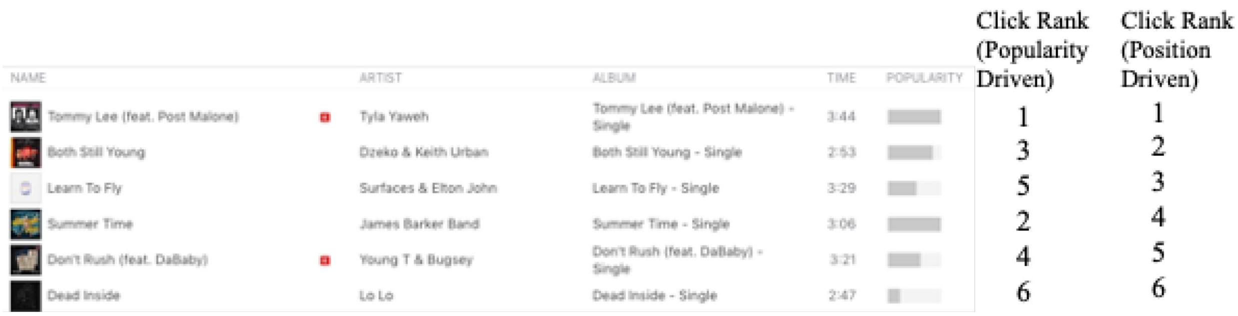
4 It is unclear to us how and why this platform does not display popularity information in a descending way. It is possible that this display is based on some recommendation algorithms. However, this is not the focus of our paper.

decisions instead of using their own judgment or even ignore their private information (Banerjee 1992, Bikhchandani et al. 1992). In the age of the Internet, popularity information has become a digital way to present other people's actions or decisions without the need to observe these actions or decisions in person. Popularity information serves as a quality signal to lower the uncertainty related to the product, especially for experience goods whose quality is difficult to obtain prior to purchase (Huang et al. 2009, Li and Wu 2018). Studies have found this popularity effect in many contexts, including music, software products, yellow pages, and e-commerce. In the music context, consumer's music consumption decision is strongly influenced by previous users' total number of listens or downloads, which is used as a proxy for the popularity information (Dewan et al. 2017, Salganik et al. 2006). In the software context, online users' choices of software products exhibit similar changes with the variations in the download ranking (Duan et al. 2009). In the context of the online yellow page, vendors receive more visits after the platform decides to display popularity information (Tucker and Zhang 2011). On the e-commerce website, positive popularity information significantly increases sales in the presence of word-of-mouth information (Chen et al. 2011).

While these studies have all shown that popularity information positively influences consumers' decisions, the effect of popularity information could be potentially overestimated or incorrectly measured. The reason is that most of the platforms that were examined in these studies have displayed products in a descending popularity order (similar to Figure 5). However, the way how popularity information is displayed on platforms like Figure 5 is consistent with the way how most people in this world read, starting from the top of the page and moves downwards. This kind of reading habit will result in the products (items) located at the top of the platform (screen) be searched or viewed more often than the products (items) located at the bottom of the platform (screen), which we refer to as the **position effect**. The product that has the largest position rank is the product located on the top row of the platform, while the product that has the smallest position rank is the one that is located at the bottom row of the platform. A classic example to show the existence of position effect is the airline reservation system, Sabre, which listed airlines based on alphabetical order. This kind of listing resulted in more bookings for American Airlines as it was listed on the top (Phillips and Journal 1988). Similar results were found in the yellow page listing, where people are more likely to view the ads on the top of the pages (Jackson and Parasuraman 1986, Lohse 1997).

Therefore, if the platform displays products in a descending popularity order (Figure 5), we will *not* be able to distinguish among whether consumer's decision is 1) driven by the popularity of the product 2) driven by the position of the product 3) driven by both effects because higher popularity also means a higher position in this situation. However, we are able to disentangle the popularity effect from the position effect if the products are not sorted by the popularity information on the platform (similar to Figure 6). On these unsorted popularity platforms, if consumers are more driven by their reading habit or the position effect, we will observe a sequential search behavior from the consumers (search from top to bottom). Whereas if consumers are more influenced by the popularity, their search pattern will be similar to the popularity rank of the product. Figure 7 visually explains how consumer's search behavior will look like for two extreme cases where consumers are either driven by the popularity information or driven by the position. The click rank number, which is a proxy for the search behavior, presents the order in which the consumer searches. A click rank of 1 suggests the product was clicked first, and click rank of 2 suggests the product was clicked second, so on and so forth. These extreme cases help us better understand how consumer's search patterns would be different, given the different driving forces. Therefore, to understand the exact effect of popularity information, we should take into account the underlying position effect by examining this kind of unsorted popularity situation. It would allow us to disentangle the popularity effect from the position effect while prior studies cannot use descending popularity platforms like Figure 5.

Figure 7 Consumer's Search Pattern for Unsorted Popularity Platform



NAME	ARTIST	ALBUM	TIME	POPULARITY	Click Rank (Popularity Driven)	Click Rank (Position Driven)
Tommy Lee (feat. Post Malone)	Tyla Yaweh	Tommy Lee (feat. Post Malone) - Single	3:44		1	1
Both Still Young	Dzeko & Keith Urban	Both Still Young - Single	2:53		3	2
Learn To Fly	Surfaces & Elton John	Learn To Fly - Single	3:29		5	3
Summer Time	James Barker Band	Summer Time - Single	3:06		2	4
Don't Rush (feat. DaBaby)	Young T & Bugsey	Don't Rush (feat. DaBaby) - Single	3:21		4	5
Dead Inside	Lo Lo	Dead Inside - Single	2:47		6	6

Aside from not being able to distinguish between the popularity effect and the position effect, prior studies are also not able to differentiate the popularity effect on different stages of consumers' decision-making. Consumers' decision process is usually considered as a two-stage decision process (Bettman 1979, Malhotra et al. 1999). In the first stage, consumers make the search decision on which options satisfy their needs for consideration. In the second stage, consumers make the purchase or consumption decision. However, prior studies have studied the influence of popularity information either only on the search decision (Dewan et al. 2017, Dewan and Ramaprasad 2012, Tucker and Zhang 2011) or only on the consumption decision (Duan et al. 2009, Li and Wu 2018, Salganik et al. 2006). Yet, several studies have shown that the same factor would have a differential impact on the two-stage' decisions. For example, a brand's perceived risk is used for the search decision. But conditioning on the search decision, the perceived risk does not have an influence on the choice decision since high-risk products are already excluded in the first stage (Erdem and Swait 2004). The position of where a hotel is listed on the Expedia website will influence consumer's click-through decision, suggesting the hotel that is listed more on the top is more likely to be clicked (Ursu 2018). But conditioning on the click decision, the position does not influence the conversion decision. As prior studies have shown the same effect could be different between the search stage and the choice stage, it is important and necessary to examine the popularity effect on both stages together.

Thus, this paper addresses the following research question: *After considering the position effect, what is the effect of popularity information on 1) consumer's search decision 2) consumer' choice decisio*n*?* The paper proceeds as follows. Section 2 outlines the hypotheses. We describe the experiment procedural and data in Section 3. Section 4 presents the empirical model specifications and results. We discuss the theoretical and managerial implications, the limitations of our study, and conclude in Section 5.

2. Hypotheses

To disentangle the popularity effect and position effect, we consider the situation where the popularity information is presented to the consumers in an unsorted way (Figure 6). Popularity information, which presents the total number of prior consumption information, is critical for consumer's decision making. Consumers tend to follow prior consumers' actions when they have limited information about the product, which is often referred to as observational learning or herding (Banerjee 1992, Bikhchandani et al. 1992). Consumers exhibit this kind of herding behavior, especially when the product's quality is unknown before purchase, such as experience products (Nelson 1970). Therefore, using popularity information can reduce the uncertainties they face (Li and Wu 2018).

Consumers are influenced by the position because it is consistent with their learned reading habits (Lohse 1997). However, the influence of popularity information should not be overridden by the position effect. Unlike the popularity information that provides the quality signal, the product's location on the platform does not provide any quality information to the consumer[5]. In other words, being on the top of the screen does not mean the product has better quality. Thus, the position of the product should not provide any additional quality signal to the consumer apart from the popularity information. Therefore, even if we account for the position effect, popularity information should still affect consumer's decisions, and this effect should be stronger than the position effect. Thus, we hypothesize that:

H1a. Popularity has an impact on the consumer's search decision even after accounting for the position effect.

H1b. The popularity effect is stronger than the position effect on the search decision.

Through the search process (first stage), consumers have gained some knowledge or information about the product. Their uncertainty level towards the product decreases after they have searched for the products. However, they still do not have full information about the product before they make the actual purchase decision. The knowledge or the information consumers obtain during the search process is still not equivalent to the full experience they will get after the purchase. In addition, some of the product qualities can only be observed after purchase (Nelson 1970, 1974). Therefore, consumers will still rely on the popularity information for their choice decision as it can further reduce the quality uncertainty through the herding mechanism.

Consumer's choice decision is still influenced by the product's position after they have made the search decision. The reason is that people are more likely to remember things that happen at the beginning of an event and also things that happen at the end of the event, which is

[5] If the platform has a recommendation system, then the location of the product has a meaning. Usually, the higher the position, the more relevant that product or the information is to the consumer. However, we do not consider recommendation systems in this study.

referred to as the primacy effect and recency effect (Miller and Campbell 1959). Studies have shown that the underlying reason for this primacy and recency effect is that the initial items presented are most effectively stored in the dormant memory while the items that are seen last are still in the active memory (Murdock and Crowder 1977). Consequently, it is easier to recall what happens first and what happens last. Therefore, when consumers are making the choice decision, they are more likely to recall the products either at the beginning of their search or at the end of this search. Thus, we will still observe consumer's choice decisions being strongly influenced by the position of the products. Therefore, we hypothesize:

H2a. Popularity information has an impact on consumer's choice decision, conditioning on search.

H2b. The position of the product has an impact on consumer's choice decision, conditioning on search.

We have hypothesized how consumers make decisions in a situation where popularity information is unsorted. Given that most of the prior studies have examined popularity information in the situation where the products have already been sorted by the popularity (Figure 5), we are interested to understand if the popularity effect is over and above the position effect in this situation. However, since we cannot cleanly disentangle popularity and position effect in this situation, we proposed an in-between situation where platforms do not display the popularity information (Figure 8).

Figure 8 Example of Platform where No Popularity Information is Presented

TITLE	ARTIST	ALBUM		
Piano Concerto No. 21 in C Major, K. 46...	Géza Anda, Cam...	Work From Hom...	2020-03-23	2:52
Kinderszenen, Op.15: 4. Bittendes Kind	Robert Schuman...	Female Pianists - ...	2020-03-23	0:59
Kinderszenen, Op.15: 1. Von fremden Län...	Robert Schuman...	Female Pianists - ...	2020-03-23	1:51
12 Etudes, Op.10: No.2 In A Minor "Chro...	Frédéric Chopin, ...	Chopin: Ballades;...	2020-03-23	1:28

Since consumers are strongly influenced by their learned spatial pattern (Lohse 1997), where they tend to search in the order of how they read (from top to bottom), we would observe a sequential search pattern when consumers are in this no popularity situation. The reason is that there is no additional information in this situation that can deviate consumers from their normal sequential search behavior. However, when popularity information is sorted in the descending order (Figure 5), it is in line with consumers' reading habits. Therefore, it makes it easier for consumers to choose the product that is popular because the search cost to find a popular product is low. In addition, because loss looms larger than gains, consumers in this situation will be reluctant to absorb the relative loss in popularity when they search down the list (Tversky and Kahneman 1991). According to reference utility theory, consumer preferences are reference-dependent in that the utility of an alternative is affected by what has already been evaluated. Thus, consumers will have less incentive to deviate from their sequential pattern compared with the situation where no popularity information is presented. Consumers will search more sequentially because the popularity is in the order with their default search behavior, and they are unwilling to bear the popularity loss. However, when popularity information is unsorted (Figure 6), based on H1b, consumers should be more driven by the popularity of the products. In this case, their sequential search behavior will deviate more from the sequential search pattern in the no popularity condition because the products' popularity is unsorted. Thus, we hypothesize that:

H3a: Consumer deviates less from the sequential search when popularity is sorted in descending order compared with when popularity information is not presented.

H3a: Consumer deviates more from the sequential search when popularity is unsorted compared with when popularity information is not presented.

3. Experimental Procedures and Data

3.1. Experimental Procedures

We conducted a randomized, between-subject experiment by using a mock website as our lab. The website mimics real-world music sites where users can listen to songs and add them to the playlist. We choose music as our research context because music is an experience good in which its quality is unknown before purchase. Therefore, consumers are more likely to use the popularity information to infer its quality. Also, music is an information good whose discovery and consumption decisions are mostly online these days. Therefore, the music context is ideal for conducting our study.

All the clickstream data from our mock music website is captured in our database. Following other scholarly works, we recruited participants on Amazon Mechanical Turk (AMT) as our experiment sample. Participants were told that they were evaluating a music website prototype and were not aware of the experiment manipulation until they were debriefed. They were given $2.5 for 25 minutes to complete the task.

Our experiment has three conditions, which are named as follows: “no popularity” condition, “popularity unsorted” condition, and “popularity sorted in descending order” condition (a.k.a. descending popularity condition). In the “no popularity” condition, participants only saw the song information along with the artists’ information. In the “descending popularity” condition, participants were presented with the songs listed in the popularity order, meaning from the most popular song to the least popular song. Daily plays are used as a proxy for the popularity information. In the “popularity unsorted” condition, the popularity information was randomly displayed along with the song list. Figure 9 presents an example of how each experimental condition looks like in our experiments.

Figure 9 Experimental Conditions

No Popularity

Song Name	Artist Name
Song 3	Artist 3
Song 2	Artist 2
Song 1	Artist 1

Popularity Unsorted

Song Name	Artist Name	Daily Plays
Song 1	Artist 1	19
Song 2	Artist 2	320
Song 3	Artist 3	25

Descending Popularity

Song Name	Artist Name	Daily Plays
Song 2	Artist 2	330
Song 3	Artist 3	27
Song 1	Artist 1	18

All the popularity information was randomly generated between 1-500. To account for the long-tail distribution in the real world, 80% of the songs are unpopular and have average popularity around 30 daily plays. 20% of the songs are popular and had average popularity around 300 daily plays. By artificially creating the gap between popular and unpopular songs, we want to strengthen their difference so that popular songs were more distinct to the user. Hence, it might increase participants' likelihood to choose it if they were driven by popularity. To control for any song-level heterogeneity, each participant saw a randomly generated list of songs. Popularity information was not associated with any songs. This means even if two participants saw the same songs, the popularity of these two songs could be different, and their position in the song list was also different. However, participants were not aware of this randomization. The songs in our database were collected from independent singers, and our pretest showed that participants were not familiar with these songs. To further control for the potential bias participants may have toward a specific singer or song name, we replace the singer name with singer X and song name with song X.

Participants were randomly assigned to one of the experimental conditions, and they were not aware of the existence of other conditions. Participants stayed in the same condition throughout the study. The experimental procedures for all studies were the same. It consisted of three main tasks, all of which were performed using our mock music website on participants' personal computers.

Task 1. Indicate Music Preference. Subjects were told that they were evaluating a prototype for an online music website. After the welcome page, subjects were asked to complete a short test of music preferences (STOMP) developed by Rentfrow and Gosling (2003) to indicate their preference for each genre on a 1-7 scale. Based on their answers, six well-known songs/music from their top two favorite genres were shown. Subjects were then asked to indicate whether they know the music and to rate them on a 1-5 scale if they knew it. The objective of these two tasks was to ensure the participants believe we were using this information to find a user who had similar music tastes as them, which would determine the total amount of extra bonus they would get. We also used the results from this task as a manipulation check for making sure the participants were knowledgeable about the genre. Participants were then told that we found a user who had exactly the same music taste as they had.

Task 2. Music Listening. After completing task 1, participants were directed to an "objective" page. They were given the instructions for completing task 2. Participants were told to choose the music they liked the most as our system had matched them with an existing user who had similar music preferences as them. They were informed that the higher similarity between participant's playlist that was created at the end of task 2 and the existing user's playlist, the higher bonus they would get at the end of the experiment. The reason for us to use this incentive is to have consumers reveal the true music preference and to be less influenced by the popularity information. However, one can argue that this incentive may lead participants to have higher chances of following other's decisions, which is to follow the popularity. If this were true, we would expect to see a higher popularity-driven behavior.

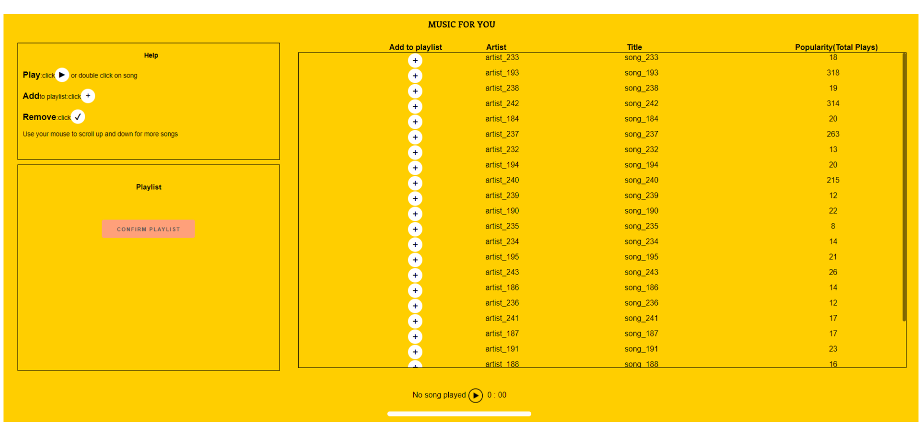
After being shown the objectives, participants then started their music listening. The music list contained 24 songs, and 12 songs were displayed on the screen so that participants had to scroll down to see the rest 12 songs. They could listen to the song by clicking its name. If they liked the music, they could add it to their playlist by clicking on the "+" button. Depending on the condition, the participants may or may not see the popularity information. Participants were required to add at least three songs to their playlist. Before they submitted their playlist, they were told to confirm that the music in their playlist was ordered by their preference.

Consumer's search decision is operationalized by their clicking behavior. We recorded the songs participants clicked on and the order they clicked on those songs. Consumer's choice decision is operationalized by whether they have added the songs they have listened to their playlist. We recorded what songs participants have added to their playlist.

Task 3. Survey. Participants completed a short exit survey that included questions on demographic information. Some of the questions were reverse coded to serve as attention check questions. They were also asked whether they have seen the popularity information and whether they remembered what it represented.

Figure 10 provides a screenshot from the mock website for the "unsorted popularity" condition.

Figure 10 Screenshot for the "Unsorted Popularity" Condition



3.2. Data

Of the 136 subjects recruited on Amazon Mechanical Turk, our final sample includes 96 subjects. We excluded the participants who failed the manipulation check (24 participants), who did not finish the experiment (9 participants), and who failed attention check (7 participants). All subjects reported living in the United States. 52% of the participants are male and have an average age of 37. 28% of the participants have a college degree, 38% have a bachelor's degree, 9% have a postgraduate degree, and others have a degree equal or lower than high school. The randomization check results suggest participants are not significantly different from each other for the three experimental conditions, suggesting the randomization was successful.

4. Analysis and Results

4.1. Descriptive Patterns in the Data

We begin the investigation by plotting a heat map of the count matrix for participants' clicking behavior in each experimental condition. The heat map is a graphical way to represent the correlation between the variable on the column and row. The column corresponds to users' click rank, and the row corresponds to either display rank or popularity rank (Table 3 shows the definition for our key variables). Each cell (click rank= i, display/popularity rank = j) represents the total number of participants who clicked on the song that has a display/popularity rank of j on their i^{th} click. In order to compare across conditions, we normalize the count matrix for each cell. Akin to a heat map, we use darker shading to indicate larger values.

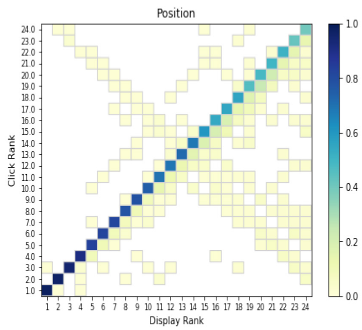
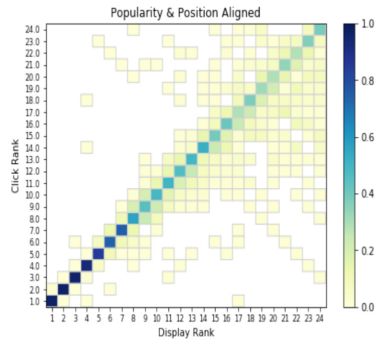
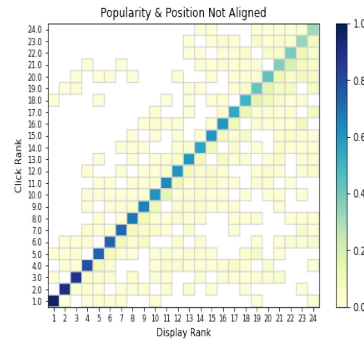
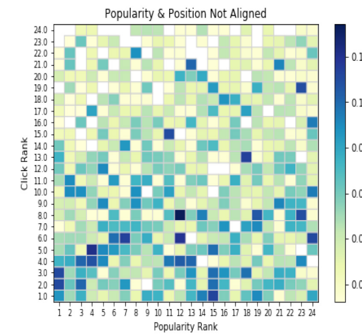
Table 3 Definition of Key Variables

Variables	Definition
Display Rank (DR)	The rank in which songs are positioned on the screen. A display rank of 1 means the song was listed on the top of the screen, and 24 means the song was listed on the bottom of the screen.
Click Rank (CR)	The rank in which individuals search the songs. Click rank equals 1 means it was the first song the subject listened to, and 2 means it was the second listened song, etc.
Popularity Rank (PR)	The rank in which the songs are ordered in their popularity. Popularity rank of 1 represents the song was the most popular, and 24 means the song was the least popular.

To examine H1, we first compare the heat maps between Table 4(c) and (d). We can see that the position effect dominates the popularity effect because the colors in Table 4(d) are very dispersed compared with Table 4(c). The popularity effect almost disappears when we randomize the display of popularity. To examine the sequential search behavior in H3a, we first compare the heat maps between Table 4(a) and (b) [6]. We can see that both heat maps have a dark shade along the diagonal line, suggesting the presence of position effect in both conditions. In Table 4(a), we see some light colors outside the diagonal line, suggesting that consumers are not searching strictly sequential when no popularity information is presented. However, the color at the diagonal line is still stronger. The same pattern exists when we compare Table 4(a) and (c). While heat maps are useful for visualization, they do not provide any statistical support for the strength of these effects. Therefore, we implement multiple empirical strategies to tackle this issue.

[6] Note that because popularity is completely correlated with position for the descending popularity condition, we are only able to plot one heat map from this condition.

Table 4 Visual Graph and Description

	No Popularity	With Popularity		
	No Popularity	Descending Popularity	Unsorted Popularity	
	Position Effect (a)	Position/Position Effect (b)	Position Effect (c)	Popularity Effect (d)
				
Average rank correlation	0.930 (0.151)	0.947 (0.104)	0.880 (0.177)	-0.047 (0.358)
Average Euclidian Distance	7.630 (10.213)	6.528 (7.389)	11.577 (9.824)	37.125 (8.867)
Average nDCG	0.919 (0.106)	0.919 (0.926)	0.859 (0.175)	0.700 (0.172)

H3a

H3b

H1a&b

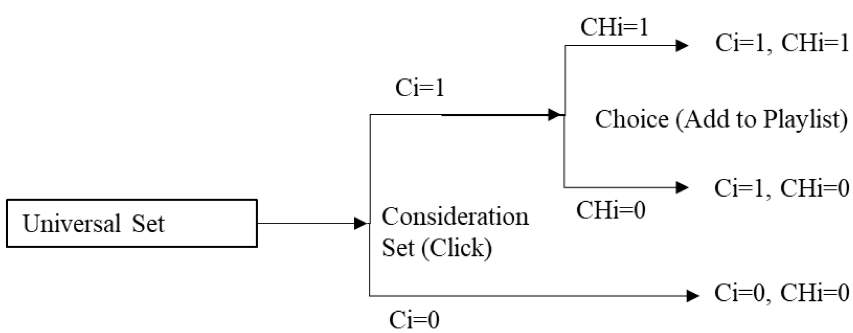
4.2. Empirical Analysis

4.2.1. Two-Stage Decision Process

We use the sequential logit model to estimate the relative effect of position and popularity on both click and choice decisions (Bettman 1979, Gensch 1987, Kardes et al. 1993).

Figure 11 presents the two-stage decision process, where the universal set represents the songs that are presented to the participant. Consumers form their consideration set by clicking the songs, and then choose songs to add to their playlist. Let C_i and CH_i denote, respectively, whether the i[th] song is clicked and chosen. C_i = 1 if the song was clicked, 0 otherwise; CH_i =1 if the song was chosen, 0 otherwise. It is worth noting that choice is only relevant when the song has been clicked at the first stage (C_i =1). Thus, if the song is not clicked (C_i = 0), the issue of choice becomes irrelevant. We follow the estimation process proposed by Kardes et al. (1993). As displayed in Figure 11, only three outcome conditions are possible: Outcome 1: C_i = 1, CH_i =1; Outcome 2: C_i = 1, CH_i =0; Outcome 1: C_i = 0, CH_i =0.

Figure 11 The Sequential Two-stage Process Model



The log-likelihood for the sequential logit model is presented as:

$$L^* = \sum \{C_i CH_i \ln[Prob.of\ outcome\ 1] + C_i(1 - CH_i)\ln[Prob.of\ outcome\ 2]$$

$$+ (1 - C_i)(1 - CH_i)\ln[Prob.of\ outcome\ 3]\}, \quad (1)$$

$where, Prob. of\ outcome\ 1 = Prob[C_i = 1, CH_i = 1],\ Prob. of\ outcome\ 2 = Prob[C_i = 1, CH_i = 0], Prob. of\ outcome\ 3 = Prob[C_i = 0, CH_i = 0]$

Let

$$F_1 = L(\beta_{0c} + X'_{ci}\beta_c) \quad (2)$$

$$F_2 = L(\beta_{0ch} + X'_{chi}\beta_{ch}) \quad (3)$$

Where (a) $L(.) = \exp(.)/[1 + \exp(.)]$, (b) X_{ci} and X_{chi} represent the vector of independent variables for the ith individual that are hypothesized to be related to the pattern of click and choice. (c) β_0, β_{0c} and β_{0ch} are vectors denoting the impact of the hypothesized variables on the click and choice stage, respectively. Therefore,

$$L^* = \sum\{C_i CH_i \ln[F_1 F_2] + C_i(1 - CH_i)\ln[F_1(1 - F_2)] + (1 - C_i)(1 - CH_i)\ln[(1 - F_1)(1 - F_2]\} \quad (4)$$

Table 5 presents the results from the estimation of the sequential logit model. We can see from Model (1) that both display rank and popularity rank have significant negative effects on the clicking behavior. This effect is negative because a popularity rank or display rank of 1 means the most popular song or the top positioned song. While examining the odds ratio, we find that the odds of a song being click decreases by 12% for display rank and 5% for popularity rank. The t-test result for the coefficients suggests that the coefficients for Display Rank and Click Rank are significantly different. Therefore, we can conclude that position has a stronger impact than popularity information on the consumer's search decision. We find support for both H1a and H1b.

Model (2) presents the second stage choice decision, which indicates consumer's choice probability after considering they have clicked on a song. Interestingly, participants' choice decision does not rely on any position or popularity information. Therefore, we do not find support for H2a and H2b where we hypothesize that the participant's choice decision is influenced by both position effect and popularity effect.

Table 5 Impact of Position and Popularity Effect on Click and Choice

	(1) Click	(2) Choice (conditioning on click)
Display Rank	-0.128***	0.031
	(0.025)	(0.030)
Popularity Rank	-0.048***	-0.029
	(0.012)	(0.015)
Click Rank		-0.051
		(0.034)
Constant	3.875***	-0.409
	(1.257)	(0.655)
Log Likelihood	-359.77	-643.07
Observations	768	560

Clustered errors in parentheses, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

4.2.2. Sequential Search Behavior

To understand the sequential search behavior, we compare the results from different experimental conditions. We compare the experimental data from "descending popularity" condition with "no popularity" condition for H3a and compare the data from the "unsorted popularity" condition with "no popularity" condition for H3b. We first use the aggregated-level data from each experimental condition to explore the results and then use the individual-level data for the analysis.

We calculate the ***Spearman's rank correlation*** ρ_j between display rank and click rank for each participant in each condition, which measures the direction of the monotonic association between these ranked variables.

$$\rho_j = 1 - \frac{6\sum_{i=1}^{n}(Display\ Rank_{ij} - Click\ Rank_{ij})^2}{n_j\left(n_j^2 - 1\right)} \quad (5)$$

Here, $Click\ Rank_{ij}$ represents the i^{th} click of the participant j, and $Display\ Rank_i$ is the corresponding display rank for that click. n_j is the total number of clicks participant j performs. ρ_j equals 1 means the subject follows a perfect top-down search pattern, ρ_j equals -1 means the subject search monotonically from bottom to top, and 0 suggests that there is no correlation between display rank and click rank. However, Spearman's rank correlation only captures how consumers are searching monotonically from top to bottom. For example, consider the situation where a user is clicking on songs that are positioned in 1, 2, 3 versus songs that are listed in 1, 13, 14, we observe the same Spearman's rank correlation.

Therefore, to capture participants' overall search behavior and to compensate for the limitations in Spearman's rank correlation, we also calculate the ***Euclidean distance*** d_j between display rank and click rank for each participant j.

$$d_j = \sqrt{\sum_{i=1}^{n}(Display\ Rank_{ij} - Click\ Rank_{ij})^2} \quad (6)$$

Again, $Click\ Rank_{ij}$ represents the i^{th} click of the participant$j$, and $Display\ Rank_{ij}$ is the corresponding display rank for that click. The smaller d_j is, the more sequentially participant j searches.

Another way of capturing participant's search behavior is to use ***normalized discounted cumulative gain (nDCG)***. Normalized discounted cumulative gain is a measure of ranking quality, which is widely used in information retrieval literature to estimate the usefulness of a document based on its position in the result list (Agichtein et al. 2006, Järvelin and Kekäläinen 2002, Wang et al. 2013). To examine the sequential search behavior, we assume the ideal search pattern is to search by the songs' sequential order, and a smaller display rank (rank equals 1 means the song is listed on the top of the list) suggests higher relevance. For simplicity, we use the reverse order of display rank to represent the relevance. The reverse of display rank of 1 is 24, the reserve of 2 is 23, and so on. Each of the reserved display rank is then divided by the log of its display rank to account for the progressive reduction in the relevance. By assigning different relevance for each search result and divided by its log value, we can better capture the top-down search because songs that are lowered in the list receive smaller relevance.

$$DCG_{DC\_ideal_j} = \sum_{i=1}^{n} \frac{Display\ Rank\ reverse_{ij}}{log_2(Display\ Rank_{ij} + 1)} \quad (7)$$

Equation 7 shows the ideal discounted cumulative gain for each participant. To obtain the actual discounted cumulative gain that portrays participants' search behavior, each participant was sorted by the click rank. Then we calculate the sum of each corresponding display rank reverse by the log of its click rank.

$$DCG_{DC\_actual_j} = \sum_{i=1}^{n} \frac{Display\ Rank\ reverse_{ij}}{log_2(Click\ Rank_{ij} + 1)} \quad (8)$$

Similarly, $Click\ Rank_{ij}$ represents the i^{th} click of the participant$j$, $Display\ Rank\ reverse_{ij}$ and $Display\ Rank_{ij}$ are the corresponding display position of that click. We then get the normalized discounted cumulative gain.

$$nDCG_{DC_j} = \frac{DCG_{DC\_actual_j}}{DCG_{DC\_ideal_j}} \quad (9)$$

$nDCG_{DC_j}$ provides insights on how much participants' sequential search behavior deviates from the ideal search behavior if they had been completely driven by position effect. If $nDCG_{DC_j}$ equals 1, it suggests that participants' search behavior is exactly the same as the perfect sequential search behavior. By assigning more weights to the top search results, accounting for the progressive decrease in rank, and normalizing against an ideal search behavior, normalized discounted cumulative gain uncovers the underlying mechanisms of how consumer searches.

We perform an OLS regression with the experiment condition indicator and demographic information (age, gender, and education).

$$DV_j = \beta_0 + \beta_1 Descending\ Popularity_j / Unsorted\ Popularity_j + \beta_2 age_j + \beta_3 male_j + \beta_4 education_j + \varepsilon_j \quad (10)$$

Here, DV_j represents Spearman's rank correlation, Euclidean distance, and normalized discounted cumulative gains for participant j. **Table** 6 (1) – (3) presents results for the association between display rank and click rank for the "descending popularity" condition and the "no popularity" condition. We hypothesize that consumers' search is more sequential in the "descending popularity" condition, suggesting that the Spearman's rank correlation between display rank and click rank should be higher for the "descending popularity" condition. This also suggests a smaller Euclidean distance and larger nDCG for "descending popularity" condition. "*Descending popularity*" is an experiment condition variable, which equals 1 if the observation comes from "descending popularity" condition and equals 0 if the observation is from "no popularity" condition. None of the coefficients for *descending popularity* are significant in Model (1) – (3), suggesting that there is no difference between the "descending popularity" condition and the "no popularity" condition in terms of display rank and click rank association [7]. Turning to the comparison between "unsorted popularity" condition and "no popularity"

[7] We use regression analysis instead of ANOVA to examine the difference across conditions as it allows us to add control variables. The downside of regression analysis is that it only shows the net differences across variables.

condition, we would expect a lower Spearman's rank correlation, larger Euclidean distance, and smaller nDCG because it is hypothesized that consumers' search will deviate more from the sequential search behavior.

Column (4) – (6) in Table 6 suggest that Euclidean distance and nDCG are significantly different for the "unsorted" condition and the "no popularity" condition. Our results can be interpreted as follows: compared with participants in the "no popularity" condition, participants in the "unsorted popularity" condition search on average 5.41 rank lower in the list, and their nDCG is 0.040 lower. This also means people in the "unsorted popularity" condition search less sequentially and deviates more from the sequential search pattern than "no popularity" condition.

Table 6 Relationship between Display Rank and Click Rank (Aggregate-level)

	Descending Popularity			Unsorted Popularity		
	(1) Rank correlation	(2) Euclidean distance	(3) nDCG	(4) Rank correlation	(5) Euclidean distance	(6) nDCG
Descending Popularity=1	0.001 (0.033)	-0.266 (2.285)	0.015 (0.010)			
Unsorted Popularity=1				-0.070 (0.042)	5.410* (2.467)	-0.040* (0.018)
Age	0.001 (0.001)	-0.021 (0.099)	-0.001 (0.001)	-0.003 (0.002)	0.196 (0.104)	-0.001 (0.001)
Male	0.019 (0.040)	-0.454 (2.778)	-0.002 (0.014)	0.023 (0.048)	-0.799 (2.736)	0.003 (0.017)
College	0.082** (0.030)	-6.103** (2.268)	0.024* (0.011)	0.086 (0.047)	-7.006* (2.800)	0.049* (0.020)
High school	0.063* (0.030)	-3.534 (2.517)	0.024 (0.012)	0.051 (0.042)	-3.380 (2.877)	0.010 (0.027)
Postgraduate	-0.061 (0.128)	2.183 (6.897)	0.018 (0.016)	-0.020 (0.113)	1.517 (6.064)	0.036 (0.022)

Constant	0.826***	10.884	0.974***	0.978***	3.380	0.999***
	(0.105)	(7.386)	(0.034)	(0.098)	(6.124)	(0.034)
Observations	64	64	64	64	64	64
R^2	0.121	0.102	0.124	0.113	0.166	0.144

Cluster robust standard errors in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Using the aggregated-level data allows us to observe the difference between each condition, but it also eliminates the individual level differences within each condition. Therefore, we use the individual-level data below to further understand the results.

To test hypotheses 3a and 3b, we estimate the below model.

$$
\begin{aligned}
Click\ Rank_{ij} = \beta_0 + \beta_1 Display\ Rank_{ij} \\
+\ \beta_2 Descending\ Popularity_j / Unsorted\ Popularity_j + \beta_3 Display\ Rank_{ij} \\
*\ Descending\ Popularity_j / Unsorted\ Popularity_j +\ \beta_4 age_j +\ \beta_5 male_j \\
+ \beta_6 education_j +\ \varepsilon_{ij} \quad (11)
\end{aligned}
$$

Table 7 presents our results. We address hypotheses 3a and 3b with Table 7(1) and (2). The interaction term in column (1) is positively significant, showing that participants in the “descending popularity” condition are more likely to click on the song that has the same display rank on an earlier click than participants in the “no popularity” condition. Turning to column (2), the coefficient for the interaction term is negatively significant. This suggests that participants in the “unsorted popularity” condition are more likely to click on the same display rank song later compared with participants in the baseline condition. Hence, hypothesis 3a and 3b are both supported.

Table 7 Relationship between Display Rank and Click Rank (Individual-Level)

	(1) Click Rank	(2) Click Rank
Display Rank	0.856***	0.857***
	(0.024)	(0.018)
Descending Popularity=1	-0.462	
	(0.255)	
Descending Popularity * Display Rank	0.067*	
	(0.029)	
Popularity Unsorted=1		0.067
		(0.338)
Popularity Unsorted * Display Rank		-0.053*
		(0.026)
Age	-0.005	-0.025*
	(0.007)	(0.010)
Male	-0.033	0.237
	(0.181)	(0.181)
College	0.366*	0.830***
	(0.151)	(0.218)
High school	0.461**	0.140
	(0.165)	(0.241)
Postgraduate	0.534	1.023***
	(0.377)	(0.284)
Constant	1.100	1.333*
	(0.602)	(0.572)
Observations	1205	1174
R^2	0.857	0.785

4.3. Additional Analysis

We conducted two additional analyses. First, we want to rule out the fact that different popularity format would play a role in the consumer decision process. Therefore, instead of displaying popularity in numbers, we use popularity bars, which look likes tally marks in Figure 6. Table 8(1) presents the two-stage model results. We find that only display rank is significant for click decision. We do not observe the popularity effect here, probably because numbers are more salient to participants compared with tally marks. Taken together, displaying popularity information in tally marks still results in a strong position effect for consumer's click decision. However, the choice decision is not affected by popularity and position effect.

Table 8 Additional Experiment on Relationship between Display Rank and Search Behavior

	(1) Click	(2) Choice (conditioning on click)	(3) Click	(4) Choice (conditioning on click)
Display Rank	-0.043***	0.012	-0.026**	0.001
	(0.011)	(0.011)	(0.008)	(0.004)
Popularity Rank	-0.008	0.005	-0.031	-0.015
	(0.010)	(0.010)	(0.033)	(0.017)
Click Rank		-0.028**		0.015
		(0.005)		(0.015)
Constant	0.521*	-0.431*	2.421***	0.026
	(0.209)	(0.183)	(.346)	(0.179)
Log Likelihood	-489.54	-695.20	-241.963	-541.46
Observations	720	338	528	538

Standard errors in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Second, we vary the search cost for participants by displaying only 12 songs. In this case, participants do not need to scroll down to discover additional songs. Still, we only observe a significant coefficient for display rank (Table 8(3)), suggesting that only position affects consumer's search process, and neither position nor popularity influences consumer's choice conditioning on click.

5. Discussion and Conclusion

By using laboratory experiments, we examine the impact of position and popularity on consumer's search and choice behavior. We find that both popularity information and the position of the product influence consumer's search decisions when the popularity information is unsorted on the platform. However, we do not find support to show that either popularity or position influence consumer's choice decision. Our results are consistent with the prior studies that examine the popularity effect on search decision (Dewan et al. 2017, Dewan and Ramaprasad 2012, Tucker and Zhang 2011). However, we show that the influence of popularity information is lower than the influence of position. This suggests that without disentangling the popularity effect from the position effect, we have overestimated the popularity influence. Our results also show that after conditioning on the search decision, popularity does not influence the choice decision, which is different from prior studies who do not consider the first-stage influence (Duan et al. 2009, Li and Wu 2018, Salganik et al. 2006). One possible explanation is that after consumers have clicked on the song and listened to it, they have formed their own opinion towards the product. Consumers will use their own judgments for the choice decision. While popularity information is a quality signal to reduce the quality uncertainty, consumers might think the uncertainty is low or almost negligible after they have searched the product. Thus, they do not rely on this information for their second-stage decision.

When comparing consumer's sequential search patterns, we find that popularity information does have an effect over and above the position effect in the "descending popularity" condition as consumers deviate less from the sequential search behavior in this situation. However, the additional popularity effect is much smaller than the position effect. This finding also implies that we cannot really distinguish the effect of popularity in this "descending popularity" condition when we are using field data. It suggests the importance of using the "unsorted popularity" situation to uncover the true influence of popularity.

To the best of our knowledge, this is the first study that examines how position and popularity affect consumer decision process through a laboratory setting, and the first attempt to understand the influence of popularity in a two-stage model. From a theoretical perspective, this paper establishes the need to disentangle the popularity effect from the position effect. Without disentangling these two effects, the "popularity effect" we observe in the "descending popularity" situation is a combination of true popularity effect and position effect. In this situation, even though the popularity information does have an effect over and above the position effect, the "popularity effect" we observe might be largely attributed by the position effect. Therefore, using unsorted popularity situation can help us disentangle these two effects and cleanly observe the behavior changes caused by popularity information. Our study also implies the need to distinguish the effect of popularity and position for search stage decision and choice stage decision. Without examining the two-stage decision process, we are attributing the influence of popularity on search decisions to the choice decision. In contrast, our study shows that popularity does not really influence the consumer's choice decision once they have experienced the product. This further implies that when platforms provide a "sampling" feature for experience goods, it can reduce the quality uncertainty associated with experience goods. By allowing consumers to "sample" the product before making consumption decisions, sampling is essentially turning experience goods to search goods.

There are also significant practical implications of our findings. First, our results suggest that platforms should leverage this position effect by charging a fee to display products on higher positions. While this strategy has already been adopted in the sponsored search, very few e-commerce websites have done that. Second, our results suggest that two-sided platforms can display the newcomers on a top position. Since position has a huge influence on consumer's decisions, if the platform keeps displaying the popular ones, the newcomers might never be discovered by the consumers. Eventually, these newcomers might leave the platform. Therefore, to ensure the balance between the two sides on the platform, the platform should consider a cold start. Third, from the product owner perspective, they should also leverage the position effect by purchasing favorable spots. Without doing so, even popular products might not be discovered by consumers.

As with any study, this work has limitations. First, our experiment uses Amazon Mechanical Turk as our research sample. One can potentially argue for the external validity by using this sample because of the relatively low payment on MTurk. Therefore, we suggest future research to take this experiment setting to field studies to generalize the findings. Second, our study only focuses on music consumption. While music can be considered as experience goods, the effect of popularity and position effect on people's consumption of search goods can potentially be different. In addition, participants can sample experience goods prior to their purchase while search goods cannot be sampled prior to purchase. Therefore, it may be a fruitful avenue for future research to replicate this study for different products. Third, our study only examines the popularity information. It would be worth examining how other information, such as online review's rating, interplays with the position, and influence consumers' two-stage behavior. Fourth, our study does not consider recommendation systems. However, using our experiment result, we can infer that consumers click on those items on the top because they have higher positions, and also consumers believe it has better suited for them. Future studies can study the situation where a highly recommended product is not shown on the top, but somewhere in between, and examine whether consumers will still choose those recommended products.

6. References

Agichtein E, Brill E, Dumais S (2006) Improving web search ranking by incorporating user behavior information. *Proc. Twenty-Ninth Annu. Int. ACM SIGIR Conf. Res. Dev. Inf. Retr.* 19–26.

Banerjee A V. (1992) A Simple Model of Herd Behavior. *Q. J. Econ.* 107(3):797–817.

Bettman JR (1979) *An Information Processing Theory of Consumer Choice*

Bikhchandani S, Hirshleifer D, Welch I (1992) A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *J. Polit. Econ.* 100(5):992–1026.

Chen Y, Wang Q, Xie J (2011) Online Social Interactions: A Natural Experiment on Word of Mouth Versus Observational Learning. *J. Mark. Res.* 48(2):238–254.

Dayton E (2020) 10 Fascinating Amazon Statistics Sellers Need To Know in 2020. *Big Commer.* Retrieved (June 18, 2020), https://www.bigcommerce.com/blog/amazon-statistics/#a-shopping-experience-beyond-compare.

Dewan S, Ho YJ, Ramaprasad J (2017) Popularity or proximity: Characterizing the nature of social influence in an online music community. *Inf. Syst. Res.* 28(1):117–136.

Dewan S, Ramaprasad J (2012) Music blogging, online sampling, and the long tail. *Inf. Syst. Res.* 23(3 PART 2):1056–1067.

Duan W, Gu B, Whinston AB (2009) Informational cascades and software adoption on the Internet: An empirical investigation. *MIS Q.* 33(1):23–48.

Erdem T, Swait J (2004) Brand credibility, brand consideration, and choice. *J. Consum. Psychol.* 31(1).

Gensch DH (1987) A Two-Stage Disaggregate Attribute Choice Model. *Mark. Sci.* 6(3):223–239.

Huang P, Lurie NH, Mitra S (2009) Searching for Experience on the Web: An Empirical Examination of Consumer Behavior for Search and Experience Goods. *J. Mark.* 73(2):55–69.

Jackson RW, Parasuraman A (1986) The Yellow Pages as An Advertising Tool for Small Businesses. *Am. J. Small Bus.* 10(4):29–36.

Jacoby J, Speller DE, Kohn CA (1974) Brand Choice Behavior as a Function of Information Load: Replication and Extension. *J. Consum. Res.* 1(1):33–43.

Järvelin K, Kekäläinen J (2002) Cumulated gain-based evaluation of IR techniques. *ACM Trans. Inf. Syst.* 20(4):422–446.

Jeffries S (2015) Why too much choice is stressing us out. *The Guardian* https://www.theguardian.com/lifeandstyle/2015/oct/21/choice-stressing-us-out-dating-partners-monopolies.

Kardes FR, Kalyanaram G, Chandrashekaran M, Dornoff RJ (1993) Brand Retrieval, Consideration Set Composition, Consumer Choice, and the Pioneering Advantage. *J. Consum. Res.* 20(1):62.

Li X, Wu L (2018) Herding and Social Media Word-of-Mouth: Evidence From Groupon. *MIS Q.* 42(4):1331.

Lohse GL (1997) Consumer eye movement patterns on yellow pages advertising. *J. Advert.* 26(1):61–73.

Malhotra NK, Peterson M, Kleiser SB (1999) Marketing research: A state-of-the-art review and directions for the twenty-first century. *J. Acad. Mark. Sci.* 27(2):160–183.

Miller N, Campbell DT (1959) Recency and primacy in persuasion as a function of the timing of speeches and measurements. *J. Abnorm. Soc. Psychol.* 59(1):1–9.

Mohsin M (2020) 10 Youtube Statistics That You Need to Know. *Oberlo*. Retrieved (June 18, 2020), https://www.oberlo.com/blog/youtube-statistics.

Murdock BB, Crowder RG (1977) Principles of Learning and Memory. *Am. J. Psychol.* 90(2):329.

Nelson P (1970) Information and Consumer Behavior. *J. Polit. Econ.* 78(2):311–329.

Nelson P (1974) Advertising as Information. *J. Polit. Econ.* 82(4):729–754.

Phillips C, Journal PT (1988) Texas Air asks transportation agency to examine rivals' reservation systems. *Wall Street Journal*

Rentfrow PJ, Gosling SD (2003) The Do Re Mi's of Everyday Life: The Structure and Personality Correlates of Music Preferences. *J. Pers. Soc. Psychol.* 84(6):1236–1256.

Salganik MJ, Dodds PS, Watts DJ (2006) Experimental study of inequality and unpredictability in an artificial cultural market. *Science (80-.).* 311(5762):854–856.

Spotify (2020) Spotify — Company Info. *Spotify*. https://newsroom.spotify.com/company-info/.

Tucker C, Zhang J (2011) How Does Popularity Information Affect Choices? A Field Experiment. *Manage. Sci.* 57(5):828–842.

Tversky A, Kahneman D (1991) Loss Aversion in Riskless Choice: A Reference-Dependent Model. *Q. J. Econ.* 106(4):1039–1061.

Ursu RM (2018) The Power of Rankings: Quantifying the Effect of Rankings on Online Consumer Search and Purchase Decisions. *Mark. Sci.* 37(4):530–552.

Wang Y, Wang L, Li Y, He D, Chen W, Liu TY (2013) A theoretical analysis of NDCG ranking measures. *Proc. 26th Annu. Conf. Learn. theory*. 25–54.

Zhu F, Zhang X (Michael) (2010) Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics. *J. Mark.* 74(2):133–148.

7. Appendix 1: Recruitment Letter

The purpose of this research is to evaluate a prototype of an online music site to be developed. As part of the research, you will be asked about your music preferences.

This study is conducted by Professor Animesh Animesh and Ph.D. candidate Qianran Jin at McGill University.

This study will last approximately 25 minutes. You will receive $2.5 plus compensation of $0-0.5 upon completion of the experiment.

Your responses will be kept confidential and will not be shared with anyone else.

If you have any concerns or questions, please feel free to contact us at any time.

Professor Animesh Animesh — Qianran Jin (student)

animesh.animesh@mcgill.ca — qianran.jin@mail.mcgill.ca

Desautels Faculty of Management

McGill University

Montreal, Canada

H3A 1G5

8. Appendix 2: Participant Consent Form

The purpose of this research is to evaluate a prototype of an online music site to be developed. As part of this research, you will first be asked to indicate your music preference by rating the songs we provide in the first stage. In the second stage, you will be asked to choose the songs you like. In the third stage, you will fill out a survey. We will match your music preference with another existing user who has the same taste in our database. The higher match of the songs you pick in the second stage with that user's choice, the higher amount of compensation you will get. You will receive a compensation of $2.5 plus $0-0.5 based on your second stage performance. The experiment takes approximately 25 minutes.

You may discontinue your participation at any time by clicking the "exit" bottom, and your responses up to this point will be discarded. If you are interrupted or have to leave the experiment and return later, you are welcome to do so. But please note that if you close your browser window, you may have to restart the whole experiment when you return.

All information that you provide will not be disclosed to anyone except the researchers; your email address will be kept confidential and will be stored on a password-protected personal computer. Other anonymous data will be kept on Amazon AWS, and only the researchers will have access to it. Individual data will not be used but will be used to aggregate all the participants. The data will be used for scientific purposes only and will be used for publication in journals.

This research is being carried out by Professor Animesh Animesh and Ph.D. candidate Qianran Jin. If you have further questions, please feel free to contact them at any time at animesh.animesh@mcgill.ca or qianran.jin@mail.mcgill.ca.

If you have any questions or concerns regarding your rights or welfare as a participant in this research study, please contact the McGill Ethics Office at (514)398-6831 or lynda.mcneil@mcgill.ca. (REB# 514-0517).

Please print a copy of this consent document for your records.

If you agree to participate in this experiment, please press the “Next” button to start the experiment.

Chapter 4. Decision Making Under Conflicting Information

1. Introduction

Consumers make decisions every day. When one alternative dominates the others, consumers can easily make the decision. For example, when choosing between one apartment that is closer to work and has better condition versus another apartment that is further to work and has a poorer condition, consumers can easily pick the first apartment because it dominates the second option. Yet, it is often more likely to experience a situation where no single alternative dominates.

Consider the following dilemmas that we often face. 1) You want to buy a TV and find two options that satisfy your requirements. One TV is 50 inches and comes with a 1080p resolution, and the other one is 43 inches but has a 4K resolution. Which one will you buy? 2) You are choosing a movie to watch for your movie night. You come across two movies. One of them has 10,000 people watched it and 900 Thumbs Up, while the other one has 1,000 people watched it but 950 Thumbs Up. Which movie will you choose?

Both examples capture the situation where there are two alternatives and two dimensions, whereas no alternative dominates. Consumers are faced with one option that is superior on dimension A but inferior on dimension B, and another option that is inferior on dimension A but superior on dimension B. However, by taking a closer look at the two examples, we can see the dimensions in these examples are different. In the first example, both of these dimensions represent the *search attribute*, meaning their qualities can be determined prior to purchase (Nelson 1970, 1974). Usually, the search attribute can be objectively quantified. Prior studies have mostly focused on examining the conflict of different search attributes on consumer's decision making (Evangelidis and Levav 2013, Fischer and Hawkins 1993, Tversky et al. 1988).

However, the dimensions in the movie example are different from the TV example in two ways. First, the two dimensions in the movie examples are interrelated. Unlike size and resolution that are independent, the number of people who have watched the movie (views) and the number of Thumbs Up are interrelated because consumers can only give Thumbs Up after they have watched the movie. This suggests that the total number of Thumbs Up is always smaller or equal to the number of people who have watched it. Second, unlike the two dimensions in the first example are both search attributes of the TV, the two dimensions in the second example are signals that proxy for the movie's experience attributes (the movie is usually considered as an experience good as it has more experience attributes than search attributes). Different from search attributes that can be objectively determined prior to purchase, experience attributes are difficult to ascertain prior to consumption (Nelson 1970, 1974). In order for consumers to assess the experience quality of a product beforehand, platforms have either subjectively quantified the experience attribute (i.e., the condition of an apartment is 90/100) or have displayed online social signals to proxy for experience attributes (Li and Wu 2018). These online social signals can include prior consumers' purchase/adoption decisions or their opinion toward the purchase, which can be used to reduce the quality uncertainties for consumers. We refer to the online social signals that represent prior consumer's purchase/adoption decisions as ***pre-consumption signals***, and the online social signals that represent consumer's opinion as ***post-consumption signals***. In our example, "the number of people who have watched the movie" is a *pre-consumption signal*. However, "the number of Thumbs Up" is not a signal for the experience attribute because it does not convey any quality information by itself. The number of

Thumbs Up is bounded by the *pre-consumption signal* as its value cannot exceed "the number of views the movie has". Therefore, it cannot reduce the quality uncertainty unless we know what is "the number of views the movie has". We refer to the information from prior consumer's opinions that cannot be used as a signal as ***post-consumption information***. However, the proportion of "the number of Thumbs Up given the number of people who have watched it" is a *post-consumption signal* as it can be used to infer the quality of the experience attributes. Previous studies have shown that consumers are likely to use either the *pre-consumption signal* or the *post-consumption signal* to make their decisions (Cai et al. 2009, Chevalier and Mayzlin 2006, Dewan and Ramaprasad 2014). However, they have not examined the influence of multiple online social signals on consumer's decision-making. Different from the search attributes and the subjectively quantified experience attributes that can be independently used for making decisions, online social signals sometimes cannot be independently used to make decisions. Imagine a scenario where video A has a 100% Thumbs Up proportion, and video B has a 90% Thumbs Up proportion. Most consumers will ask how many people have viewed these videos before making their choices. What if the 100% Thumbs Up proportion is based on ten viewers, and the 90% Thumbs Up proportion is based on 1000 viewers? This example suggests that using either *pre-consumption signals* or *post-consumption signals* does not provide a complete picture for consumers to make decisions. While consumers can use one of the signals for making decisions if the other one signal is absent, it only provides partial information to the consumers.

Given the two differences mentioned above, it is important and necessary to examine the influence of online social signals together on consumer's decision making. The multi-attribute value model has been widely used to understand how consumers make choices when there is more than one product attribute (Keeney and Raiffa 1979), where the influence of all attributes is the weighted sum of each attribute's value weighted by the attribute weight. However, this model is not suitable for our context because the model implicitly assumes the attributes have to be independent. However, the online social signals in our context are interrelated. Therefore, it remains unclear how consumers make decisions when they are given multiple online social signals, which are proxy for the experience attributes of the products. Thus, it is critical for us to examine the following research question: *How do consumers make decisions in the presence of multiple online social signals?*

We conducted three experiments to answer this research question. Our results suggest that consumers are bounded rational. While they think *post-consumption signals* are more important than *pre-consumption signals*, their choices do not follow this perspective. When one alternative is dominant on *pre-consumption signal* and post-consumption information, consumers are very likely to disregard the underlying *post-consumption signal* and simply prefer the option with the higher face value. However, when *pre-consumption signals* and post-consumption information are conflicting, consumers will then take into account the *post-consumption signal* for making decisions. The consumer's decision strategy largely depends on the relationship of *pre-consumption signal* and *post-consumption signal*. When the *post-consumption signal* is explicitly presented, more consumers will use this signal and flip their choices.

Our contributions are three-folded. First, we extend the current studies on examining the influence of multi-attributes on consumer decision making into understanding the impact of multiple online social signals by proposing a Bayesian inference approach. Second, we propose a decision heuristics to explain people's choices when facing different social signals. Third, our study brings together the two online social signals that were studied separately in previous studies, *pre-consumption signal* and *post-consumption signal*, and examines their influence in the same context. Our finding makes critical theoretical contributions and significant practical implications in guiding how platforms should display the information.

The paper proceeds as follows. Section 2 provides a brief overview of the related literature. Section 3 discusses two decision strategies as the theoretical background. Section 4 to Section 6 describes the three experiments and the results. We conclude and discuss our contributions, managerial implications in the last section.

2. Related Literature

2.1. Multi-attribute Decision Making

Our decisions making usually involves comparing more than one attribute of the product. One stream of research in studying how consumers make decisions when facing multi-attribute products have assumed that consumers are rational decision-maker (Fishbein 1976, Keeney and Raiffa 1979, Rosenberg 1956). Consumers have well-defined preferences for the option, which do not depend on how the options are described and how preferences are elicited. Consumers evaluate the attributes for each alternative and calculate a weighted sum for all attributes. The option that has the largest sum will then be chosen. This kind of weighted additive rule is often seen as a complex algorithm that results in optimal decisions and accurate judgment (Payne et al. 1993). The attributes are implicitly assumed to be independent of each other, while each attribute's value and weight can contain uncertainty (Kahn and Meyer 1991, Meyer 1981).

Another stream of research has argued that consumers do not really behave like the rational choice model because they are bounded rational and have limited processing capacity (Simon 1955). Rather than having a well-defined preference on the choices, consumers' preferences are constructed on the spot when needed. Decisions can depend on multiple factors, including the context or how other options are presented (Tversky and Simonson 1993), the way preferences are elicited (Tversky et al. 1988), and how the options are framed (Kahneman et al. 1981).

Both streams of research provide insights on how consumers make decisions, especially when there is a conflict. However, the attributes that are studied in both streams are the inherent attributes of the product. While in our context, the "attributes" we examine are the online social signals that are used as a proxy for the product's inherent attribute. Therefore, it is unclear from the prior research which perspective would work better in explaining consumers' choice in this context. Thus, we use both rational and bounded rational perspective to understand consumer's decision making in this study.

2.2. Search Attribute, Experience Attribute, and Online Social Signals

Product attributes can generally be classified into two types, search attributes and experience attributes (Nelson 1970, 1974). The difference lies in how difficult it is to obtain the attribute's quality before consumption. Consumers can obtain full information of search attributes prior to purchase, while the information of experience attributes can only be known after purchase or after using the product. To reduce the difficulty or cost of obtaining the information on experience attribute prior to purchase, platforms start to display information from prior consumer's purchase and opinion. This information serves as a signal to reduce the uncertainty in experience attribute prior to consumer's purchases (Klein 1998). Two types of signals have been studied in past literature, prior consumer's purchase/adoption information, which is referred to as ***pre-consumption signals*** or *observational learning*, and prior consumers' opinion towards their purchase decisions, which is the ***post-consumption signals*** or *word-of-mouth (WOM)*.

Consumers tend to follow what others are doing rather than using their own information when they have limited information (Banerjee 1992, Bikhchandani et al. 1992). The process of which they make decisions based on prior users' aggregated decisions is called observational learning (OL). The influence of OL information or *pre-consumption signals* on subsequent consumer's decision making has been examined in multiple contexts. In the music context, the total number of "listens" a song has by previous users, which is a proxy of *pre-consumption signal*, positively influences consumers' subsequent listening decisions (Dewan and Ramaprasad 2012). In the context of the software product, online users' choices of software products exhibit distinct jumps and drops with changes in download ranking (Duan et al. 2009). In the context of online yellow pages, online niche vendor receives more visits than popular vendor after popularity information is presented (Tucker and Zhang 2011).

The research on *post-consumption signals* has mostly examined the valence and volume of word-of-mouth (WOM) on product sales and consumer decision making. While WOM volume positively influences sales (Chevalier and Mayzlin 2006), the relationship between WOM valence and sales are mixed. Some studies have shown that review rating positively influences sales (Chevalier and Mayzlin 2006), while others have suggested that rating has no significant impact on sales (Duan et al. 2008, Liu 2006). In terms of the WOM valence generated by the crowd and friend, there is a herding behavior of subsequent rating for movies with a large volume of crowd rating and a differentiation behavior for movies with a smaller volume of rating. Moreover, the friend's rating has a herding effect. Similarly, the total number of likes from both crowd and friends affect people's decision (Dewan et al. 2017).

While *pre-consumption signals* and *post-consumption signals* independently influence consumer's decision making by being the proxy of experience attributes, little is known about their joint influence on the decisions. In addition, prior studies have not examined consumer's relative strength on these two signals. We argue that since *pre-consumption signals* only represent prior consumer's actions, it is a weaker signal compared with *post-consumption signals*. Thus, consumers will rely more on the stronger signal, *post-consumption signals*, for their decision making.

2.3. Online Social Signals in Our Study

Pre-consumption signals and post-consumption information are examined in this study. Specifically, we use "the total number of views (View)" of a product as the *pre-consumption signals*, and "the total number of Thumbs Up" (Thumbs Up) a product has as the post-consumption information. Although previous studies have mostly studied post-consumption information using a 1-5-star rating (Chevalier and Mayzlin 2006, Mudambi and Schuff 2010), the star rating is more complicated than the simplified Thumbs Up as it involves five different levels of opinion (Balboni 2020). Therefore, using Thumbs Up can reduce consumer's cognitive effort in making decisions.[8] In addition, major video platforms, such as YouTube and Netflix, have changed from star ratings to Thumbs Up these days, which confirms the validity of using this measurement.

Different from search attributes that can be objectively determined and are usually independent, View and Thumbs Up are intercorrelated because people can only give out a Thumbs Up after he/she has viewed the product. Thus, Thumbs Up is a subset of View. When seeing the subset relationship, it is natural for people to construct the proportion, $\frac{Thumbs\ Up}{View}$. The

[8] We acknowledge that it is our study's limitation for not considering 1-5-star rating. Therefore, we strongly suggest future research to look into the different effect of star rating and thumbs up. For the time being, we also do not explicitly consider Thumbs down for our study.

proportion represents the percentage of people who give Thumbs Up out of the total population who have viewed it and is a representation of *post-consumption signals.* We assume that consumers will rely more on the $\frac{Thumbs\ Up}{View}$ proportion than View or Thumbs Up to make their choices as this proportion has taken into account both View and Thumbs Up. Therefore, it is more informative than the *pre-consumption signals*, View, and Thumbs Up, which is only a piece of information, not a signal.

One thing to notice here is that making decisions solely on View or $\frac{Thumbs\ Up}{View}$ proportion is insufficient. Suppose we have Option A that has “5 View and 5 Thumbs Up”, which is 100% $\frac{Thumbs\ Up}{View}$ proportion and Option B has “1000 View and 900 Thumbs Up”, which is also 90% Thumbs Up Like/View ratio. If we only use the *pre-consumption signal* View, then Option B will be preferred. If we only look at the *post-consumption signal* proportion, Option A is more likely to be chosen. Therefore, we need to consider both signals when making the decision, and the Bayesian view can take into account that.

3. Theoretical Background

3.1. Bayesian Inference Decision Strategy

Before discussing the Bayesian Inference, we should dig deeper into the two online social signals we are studying here. While View is the number of consumers who have viewed the product, $\frac{Thumbs\ Up}{View}$ is the proportion of consumers who give Thumbs Up after viewing the product. The underlying process of how the value of these two social signals is generated is very similar to a binomial distribution. Each consumer can be considered as an individual experiment, and View represents the total experimenters or samples we have. $\frac{Thumbs\ Up}{View}$ proportional is the number of success proportion, in which "success" is the analogy for a Thumbs Up.[9] Therefore, each alternative follows a binomial distribution B (View, $\frac{Thumbs\ Up}{View}$). A Bayesian view of this problem can naturally take into consideration the View to "correct" for $\frac{Thumbs\ Up}{View}$ proportion. Assume we have a Beta distribution (1,1), a common conjugate prior for the binomial distribution, which means the probability of seeing all proportions are equally likely. Then the posterior estimates or the probability of observing the given Thumbs Up/View proportion in the whole population will follow a beta distribution (Thumbs Up +1, View – Thumbs Up +1). For two options A and B, we can calculate the probability that Option A's $\frac{Thumbs\ Up}{View}$ proportion is greater than Option B's $\frac{Thumbs\ Up}{View}$ proportion in the whole population. If this probability is greater than 0.5 (out of 1), suggesting that the proportion of Option A is more likely to be greater than Option B's proportion, then Option A should be more preferred. Otherwise, Option B will be chosen. Mathematically, this probability is calculated as below.

$$p = \frac{Thumbs\ Up}{View}$$

$$\alpha = Thumbs\ Up + 1$$

$$\beta = View - Thumbs\ Up + 1$$

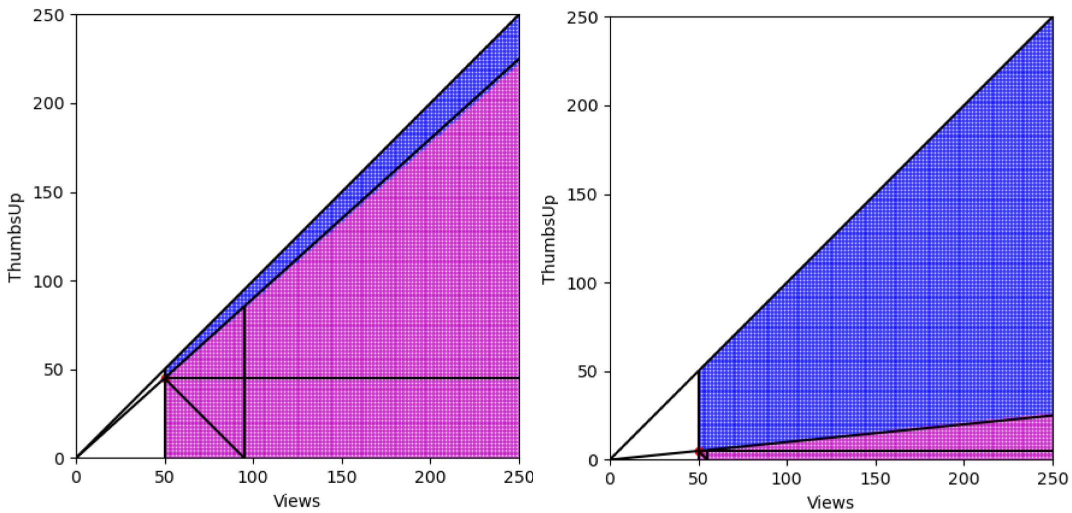
$$\Pr(p_A > p_B) = \sum_{i=0}^{\alpha_A - 1} \frac{B(\alpha_B + i, \beta_A + \beta_B)}{(\beta_A + i)B(1 + i, \beta_A)B(\alpha_B, \beta_B)}$$

Based on this equation, we simulated two potential scenarios to predict how consumers will choose when we fixed the values for Option A (Figure 12). The blue area indicates

9 Here, we assume that if consumer does not give a Thumbs Up, it is counted as a "failure". While in reality, consumers who do not give Thumbs Up can either be people who want to give Thumbs Down or people who are neutral to the product. We will discuss the potential bias of this assumption later in the paper.

consumers will prefer to choose Option B, whereas the pink area suggests consumers will prefer A.

Figure 12 Bayesian Inference Graph



3.2. Three-Stage Decision Strategy

While the Bayesian Inference approach can be used by rational decision-makers, we also discuss the decision strategy that bounded rational consumers might use. To further examine this decision strategy, we categorize three regions in which the alternatives lie in and discuss how each region is different. Figure 12 shows the three categories in a different color. We assume Option A is fixed for easier illustration.[10] We also assume Option A's View is always smaller than Option B's View because of symmetricity. The line that crosses origin point and Option A has the same slope as Option A's $\frac{Thumbs\ Up}{View}$ proportion. Therefore, any points that are located above this line has a larger $\frac{Thumbs\ Up}{View}$ proportion than Option A and vice versa. Below, we discuss the three categories based on the regions and summarize them in Table 9

[10] Note that the value of Option A would not influence the three regions we describe. It only influences the area each region covers. But the relationship between Option A and Option B in each region will not change.

.1) When Option B is positioned in the green area (Region 1), it suggests that Option B's View and Thumbs Up are both larger than Option A. In addition, its $\frac{Thumbs\ Up}{View}$ proportion is also greater than Option A. We call this region a "*completely dominant region*".

2) When Option B takes the value in the pink area (Region 2), it suggests that Option B's View and Thumbs Up are both larger than Option A. However, unlike the first category, Option B's $\frac{Thumbs\ Up}{View}$ is smaller than Option A. We refer to this region as a "*seemingly dominant region*" because only by looking at the View and Thumbs Up, Option B has larger values than Option A.

3) When Option B takes the value in the yellow area (Region 3), it suggests that Option B's View is larger than Option A's View. However, Option B's Thumbs Up and $\frac{Thumbs\ Up}{View}$ are both smaller than Option A. We refer to this region as a "*conflicting region*".

Figure 13 Region Plots

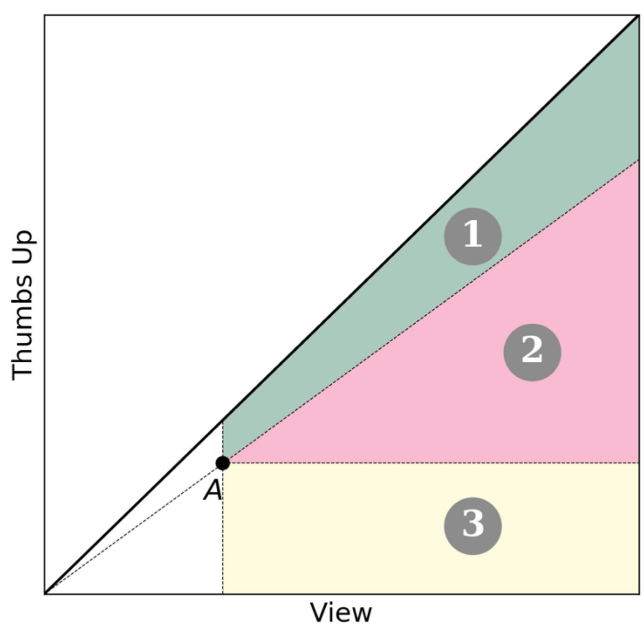


Table 9 Decision Categories

Category	Relationship
Completely Dominant (Region 1)	$View_A < View_B \quad Thumbs\ Up_A < Thumbs\ Up_B$ $\left(\frac{Thumbs\ Up}{View}\right)_A < \left(\frac{Thumbs\ Up}{View}\right)_B$
Seemingly Dominant (Region 2)	$View_A < View_B \quad Thumbs\ Up_A < Thumbs\ Up_B$ $\left(\frac{Thumbs\ Up}{View}\right)_A > \left(\frac{Thumbs\ Up}{View}\right)_B$
Conflicting (Region 3)	$View_A < View_B \quad Thumbs\ Up_A > Thumbs\ Up_B$ $\left(\frac{Thumbs\ Up}{View}\right)_A > \left(\frac{Thumbs\ Up}{View}\right)_B$

Based on this categorization, we discuss the three-stage decision strategies proposed by Tversky et al. (1988) and Fischer and Hawkins (1993). They suggest that people make decisions in three steps in a two-alternative two-attribute scenario, i.e., Option A and Option B. First, people examine whether Option A dominates Option B in all perspectives, or equal in most perspectives but dominates in at least one. If there exists such dominance, then Option A will be preferred. And vice versa. If no dominance relationship exists, then in the second step, people examine whether one option has a decisive advantage over the other alternative. In other words, whether the difference between the two alternatives on one attribute is much greater than the difference on the other attributes. If one option has such decisive advantage, it will be chosen in this step. Third, if the options do not have decisive advantage, people will follow a lexicographic procedure and choose the product that is superior on the more important attribute. If this is the underlying heuristic for consumer's choices, we would observe people favoring Option B in the *completely dominant region* and the *seemingly dominant region*. People will prefer Option A in part of the *conflicting region* because $\frac{Thumbs\ Up}{View}$ proportion is more important than View.

However, they will prefer Option B in the right-hand side of this region (when Option B is further away from Option A) because Option B has a decisive advantage on View.

Below we summarize the expected decisions under Bayesian Inference and three-stage decision strategy.

Table 10 Expected Decisions

	Region 1	Region 2	Region 3
Bayesian Inference	B	A	A
Three-stage Decision Strategy	B	B	A B (when B's view has a decisive advantage)

We can see that the difference between these two decision strategies mainly lies in region two and region three. Therefore, to better understand which decision strategy consumers follow when making the decisions, we conduct two experimental studies focusing on the *seemingly dominant region* and *conflicting region*. While consumer's decision in the completely dominant is also very interesting, it is not the focus of this paper.

4. Study 1: Seemingly Dominant Region

The core object of study 1 was to examine how consumers make decisions in the *seemingly dominant region*. Thus, we constructed pairs of options where one option's View and Thumbs Up values are greater than the other option while its $\frac{Thumbs\ Up}{View}$ proportion is smaller. We tested which option consumers prefer when facing different alternatives.

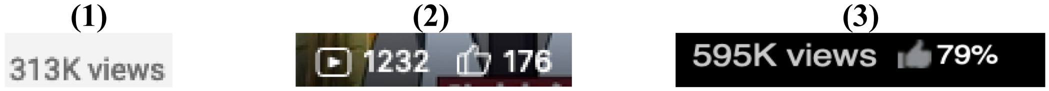
4.1. Method

4.1.1. Study Context

We choose online video as the study context for three reasons. First, online videos are experience product that has relatively more experience attributes than search attributes. Therefore, consumers tend to rely on signals for experience products to infer the video's actual quality beforehand. Second, online videos have both *pre-consumption signals* and post-consumption information. In order to express their opinion on the video by giving Thumbs Up, consumers technically have to first watch the video. Third, online videos have multiple ways of displaying these signals. Figure 14 shows the screenshots of how three different platforms display their videos. On the first platform, information about how many people have viewed the video (*pre-consumption signal*) is displayed. On the second platform, the view information (*pre-consumption signal*) and the number of people who thumbs up on the video (post-consumption information) are shown. On the third platform, the number of views (*pre-consumption signal*) is displayed along with the Thumbs Up proportion (*post-consumption signal*). With so many different ways of displaying the information, it becomes critical to understand what works best for consumers to make decisions. Our study mainly uses the second platform's display, which has View and Thumbs Up information, because it includes the information that is being displayed on platforms 1 and 3.

Figure 14 Example of Information Displayed on Three Platforms

(1) **(2)** **(3)**



4.1.2. Procedure

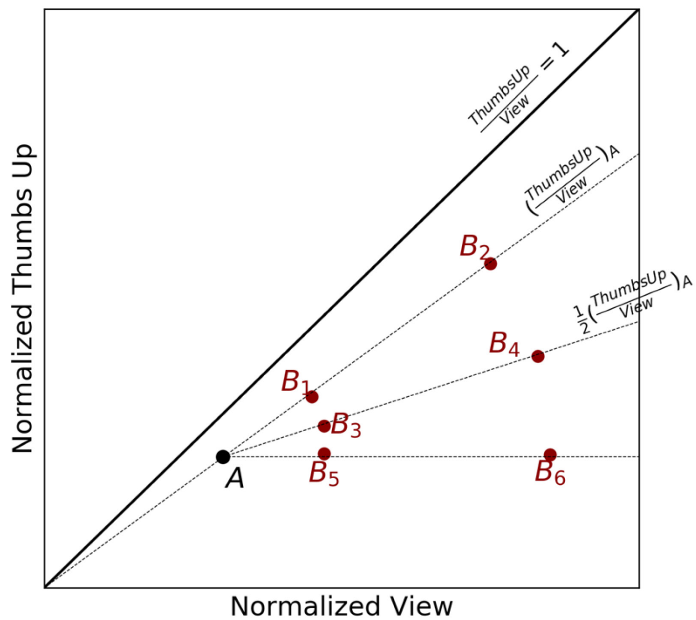
One hundred and twenty participants (55% male, M_{age}=40 years) were recruited through Amazon Mechanical Turk and were asked to participate in a survey on "understanding decisions". Participants are randomly assigned to the experimental conditions. They were asked to "imagine that [they are] consider watching a video" and were asked to choose one of two the videos presented to them in each choice sets. Each participant complete six video choices. The study concluded with a basic demographic questionnaire.

4.1.3. Choice Sets

For each of the six pairs, Option A's View and $\frac{Thumbs\ Up}{View}$ proportion are kept constant on the normalized value for the participant throughout the experiment. To illustrate how we generate Option B, we plot the choice sets in Figure 15. For a given Option A, we generate six alternative Option B, which is represented by Option B_1-B_6. Similar to Option A, Option B_1-B_6 are also presented in normalized values in the figure.

We choose Option B_1 and Option B_2 in which their $\frac{Thumbs\ Up}{View}$ proportions are slightly smaller than Option A's $\frac{Thumbs\ Up}{View}$ proportion. Option B_3 and Option B_4 are chosen on the line where its slope is half of Option A's $\frac{Thumbs\ Up}{View}$ proportion. Note that this does not mean Option B_3 and Option B_4's $\frac{Thumbs\ Up}{View}$ proportions are half of Option A's $\frac{Thumbs\ Up}{View}$ proportion. Option B_5 and Option B_6 are constructed with having slightly larger Thumbs Up compared with Option A.

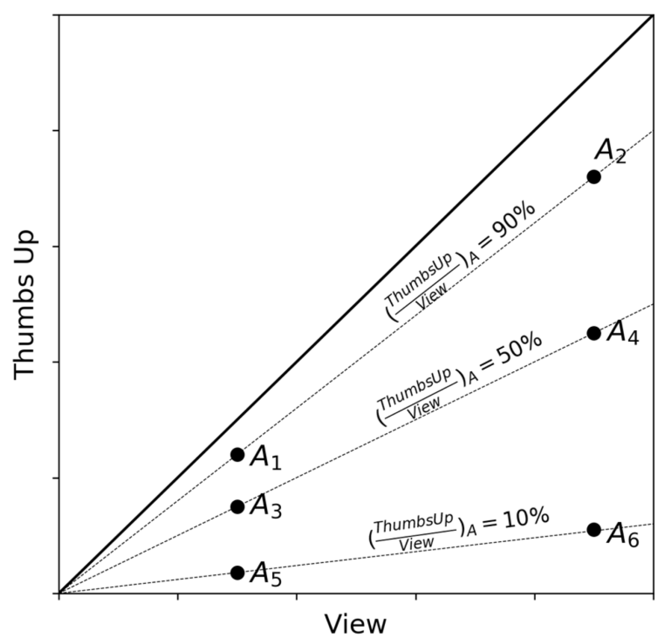
Figure 15 Example of Choice Sets in Study 1



To ensure the consistency of the scaling of Option B across the experimental conditions, B_1, B_3 and B_5 are chosen on the same contour line, suggesting the Euclidean distance between these Option B and Option A is the same. Similarly, B_2, B_4, B_6 are chosen on the same contour line. The distance between each B_1, B_3, B_5, and A is smaller than the distance between each B_2, B_4, B_6 and A. With these normalized values of Option B, we construct six choices sets using Option A' View value that is mentioned below.

While the main interest of this study is to examine consumer's choices given the different relationships between Option A and Option B, we assign different View values to Option A. In this way, the options in the experiment are more similar to the real-world videos where View counts can vary. Option A's View count is constructed by having low and high magnitude with three different levels of $\frac{Thumbs\ Up}{View}$ proportion (10%, 50%, 90%). A low magnitude of View takes a random number from 100 to 500, while a high magnitude of View takes a random number from 1 million to 2 million. We choose these numbers based on the view counts on YouTube (Cheng et al. 2008), which follows a long-tail distribution. Low magnitude represents the tail of the view count, and high magnitude represents the head of the view count. When displaying the View count for high magnitude condition, we use “M” to represent million rather than writing out the whole number (i.e., 1.2 M instead of 1,201,186). Our pre-test results suggest that when displaying numbers in all seven digits, participants are more likely to randomly make a choice, thus creating more noise to the data. In addition, most video platforms also use K or M to present thousand or million. $\frac{Thumbs\ Up}{View}$ proportion determines the Thumbs Up value for Option A. For example, if Option A is low View magnitude and 50% $\frac{Thumbs\ Up}{View}$ proportion, then it will have a View value ranging from 100 to 500 counts, and a Thumbs Up value that equals to the View count times 50%. We visually plot choices of Option A’s View value in Figure 16.

Figure 16 Option A's Choices in Study 1



For each Option B_1-B_6 in Figure 15, even though Option A’s normalized value is kept constant, the actual value for Option A in these six pairs are different because of randomization. Thus, with these six versions of Option A’s View and the six corresponding Option B for each given Option A, we have 36 pairs of choice sets in total. Since having participants complete 36 choices would be too much, we randomly assign the participants into one version of Option A. Participants were told to assume that the two alternatives in each choice set were similar on all other aspects except for the View and Thumbs Up. These six choices are presented in a random order to reduce any carry-over effect.

4.2. Results

We first use a scatterplot to display our results in Figure 17. We can see that the percentage of participants choosing Option B is different depending on its location with Option A.

Figure 17 Share of Option B in Study 1

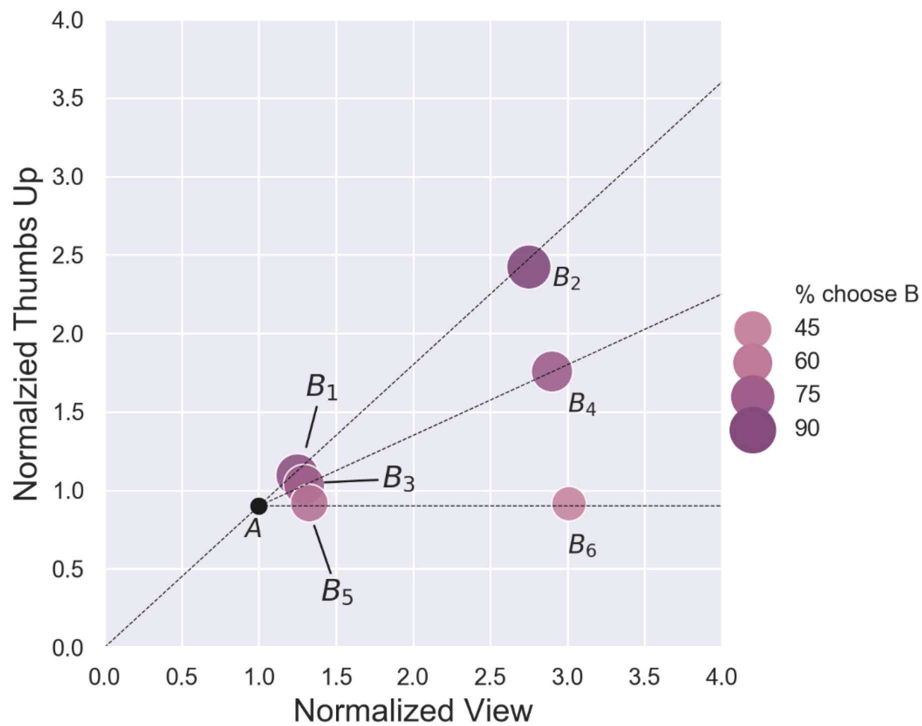


Table 11 displays the percentage of participants choosing Option A and Option B. While Option B is preferred over Option A in all six choice pairs, the percentage is different (χ^2 =

48.17, p<0.001). When Option B has a similar $\frac{Thumbs\ Up}{View}$ proportion as Option A, it has the largest percentage of participants preferring Option B (83%, 88%). When Option B's Thumbs Up value is similar to Option A, the percentage of people preferring Option B over Option A is the smallest (64%, 54%).

Table 11 Results from Study 1

	B_1	B_2	B_3	B_4	B_5	B_6
Choose A	18%	13%	25%	24%	37%	47%
Choose B	83%	88%	75%	76%	64%	54%

4.3. Discussion

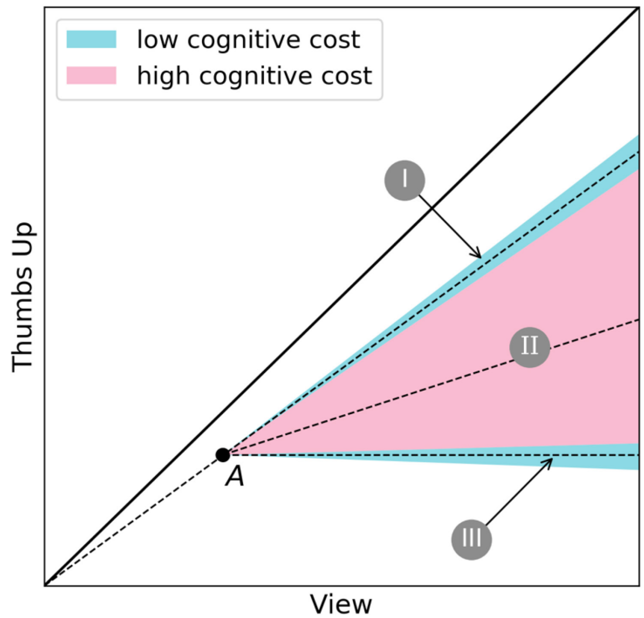
Our results in the *seemingly dominant region* suggest that consumers do not behave the same way as Bayesian Inference strategy since it predicts consumers will always prefer Option A over Option B. Our results seem similar to the three-stage decision strategy since Option B is more preferred in this region. However, it is different. While the three-stage decision strategy deterministically predicts people will prefer B in this region, our findings show that the percentage of people preferring Option B will change depending on Option B's relationship to Option A. In addition, our findings indirectly provide the support that some consumers do take into account $\frac{Thumbs\ Up}{View}$ proportion when making decisions. If they completely ignore this *post-consumption signal*, we should observe the same or similar percentage of consumers choosing Option B in all choice pairs because Option B always dominates Option A on View and Thumbs Up. In addition, if consumers only consider View and Thumbs Up to make the decisions, they are more likely to choose Option B when it is located on the larger Euclidean distance line (B_2, B_4, B_6) than the smaller Euclidean distance line (B_1, B_3, B_5). However, our results show that the percentage of choosing Option B is smaller on B_6 (54%) than B_5 (64%). This clearly suggests that people take $\frac{Thumbs\ Up}{View}$ proportion into account when making decisions because B_6's $\frac{Thumbs\ Up}{View}$ proportion is smaller than B_5's $\frac{Thumbs\ Up}{View}$ proportion.

We propose that consumers can be categorized into three groups in our study context, depending on their decision strategy. One type of consumer will only examine the **face value** of the options. In this case, they will only consider the information we provided to them, which is View and Thumbs Up. They will make the decisions based on which dimension they think is more prominent, either View or Thumbs Up. The decision strategy of the second type of consumer is based on the **minimum cognitive load**. They always go with the choices that require the least effort. The third type of consumer uses **signals** for their decision making. Instead of the face value, they look for *pre-consumption signals* and *post-consumption signals*. Since the signals, especially *post-consumption signals*, are not as easy to obtain as face values, they might be bounded by the cognitive cost.

In Figure 18, we argue that the cognitive cost to examine the difference between Option A and Option B's *post-consumption signal*, $\frac{Thumbs\ Up}{View}$ proportion, is smaller when Option B is in the "similar $\frac{Thumbs\ Up}{View}$ proportion region" (region I) or the "similar Thumbs Up region" (region III) compared with when Option B is in region II. For consumers who use signals to make their choices and are cognitively bounded, they can easily infer that the $\frac{Thumbs\ Up}{View}$ proportion is similar for Option A and Option B in region I. Therefore, their decisions are based on the option that dominates *pre-consumption signals*. In "similar Thumbs Up region" (region III), the cognitive load to identify the difference between Option A and Option B's $\frac{Thumbs\ Up}{View}$ proportion is also low. People can tell that Option B's *post-consumption signal* is much smaller than Option A. Therefore, they will prefer Option A. Because they are cognitively bounded, their decision in

Region III depends on which alternative has larger *pre-consumption signal*, which is Option B in this case.

Figure 18 Cognitive Cost for Making Decisions



For people who are cognitively capable of examining the $\frac{Thumbs\ Up}{View}$ proportion, they will prefer the option that dominates on the *post-consumption signal*. However, some of them might think the advantage of Option A's *post-consumption signal* is incomparable compared with the huge advantage of Option B' *pre-consumption signal* in region I. In this case, Option B will be more preferred for them. Table 12 summarizes the different types of potential consumers in each region.

Table 12 Consumer Decision Heuristics in Seemingly Dominant Region

	Cognitive Load	Region I	Region II	Region III	Decision Heuristic
Face Value		B	B	B	Choose the option with a larger View if View is more important.
		B	B	B	Choose the option with larger Thumbs Up if Thumbs Up is more important.
Minimum Cognitive Load		B	B	B	Use the minimum effort to make decisions. Therefore, they use face value for this region.
Signals	Cognitively bounded	B	B	A	For regions that require less cognitive load and can easily tell the options are similar, choose the option with a larger *pre-consumption signal*. For regions that require less cognitive load, and can easily tell the options are different, choose the option with a larger *post-consumption signal.* Otherwise, choose the option with a larger *pre-consumption signal.*
	Cognitively capable	B	A	A	Choose the option with a larger *pre-consumption signal* when the *post-consumption signal* is similar. Otherwise, choose the option with a larger *post-consumption signal*.
		A	A	A	Choose the option with a larger *post-consumption signal*.

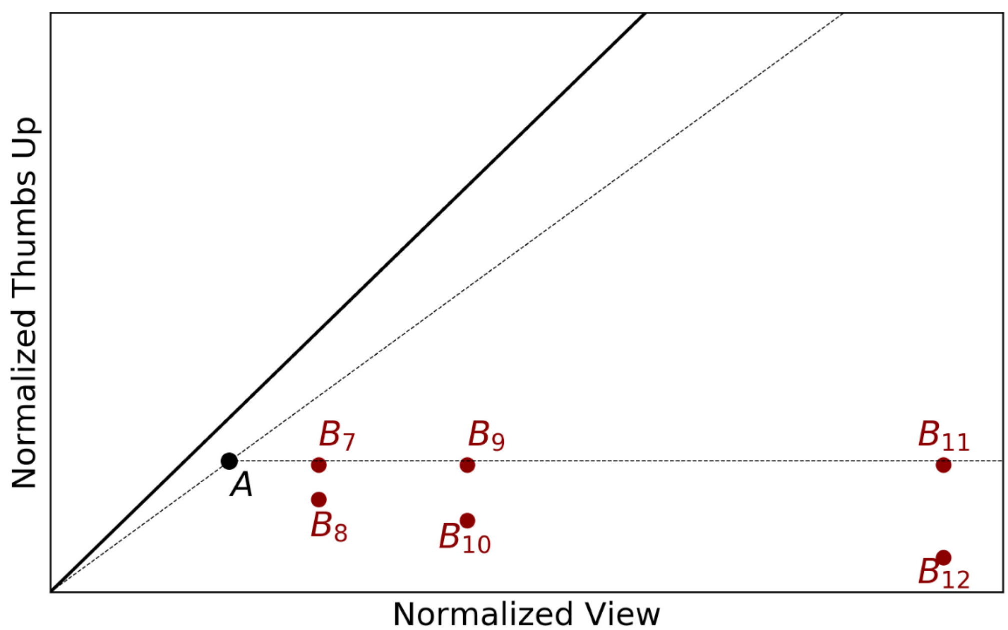
5. Study 2: Conflicting Region

The core object of study 2 is to examine how consumers make decisions in the *conflicting region*. Thus, we constructed pairs of options where one option has a larger value of View, and the other option has a larger value of Thumbs Up. Here, the option that has larger Thumbs Up value also means it has a larger $\frac{Thumbs\ Up}{View}$ proportion by construction.

5.1. Method

Eighty participants (46% male, M_{age}=39 years) were recruited through Amazon Mechanical Turk and were asked to participate in a survey on "understanding decisions". The experiment procedure is the same as Study 1. Participants completed six video choices and a basic demographic questionnaire.

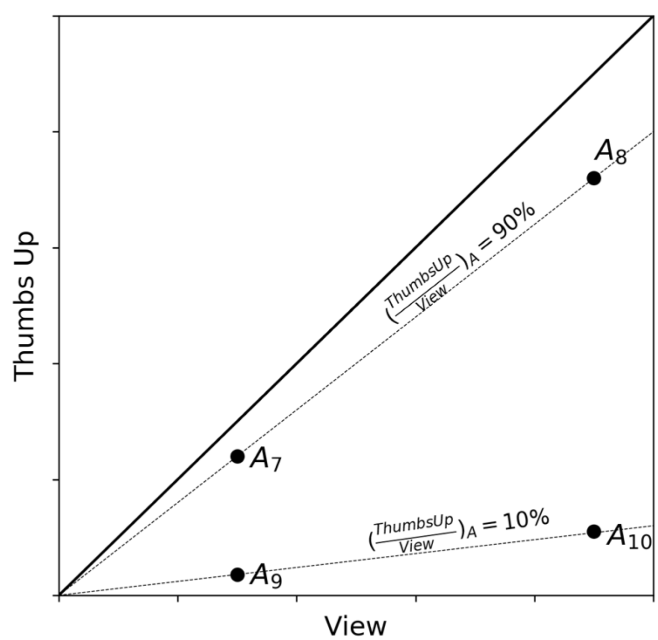
Figure 19 Example of Choice Sets in Study 1



Similar to Study 1, Option A's View and $\frac{Thumbs\ Up}{View}$ proportion are kept constant on the normalized value for the participant throughout the experiment. Figure 19 presents the choice set for this study using normalized values. Option B_7, B_9 and B_{11} are constructed with having slightly smaller Thumbs Up value compared with Option A. Option B_8 and B_{10}'s Thumbs Up values are about half of Option A's Thumbs Up value, and Option B_{12}'s Thumbs Up value is about a quarter of Option A's Thumbs Up value. For easier comparison of Study 2's results with Study 1, Option B_7, B_8 have the same View value as Option B_5 in Study 1 and Option B_9, B_{10} have the same View value as Option B_6 in Study 1. Option B_{11}, B_{12} have a higher magnitude of View compared with Option A (11 times).

Similar to Study 1, we also assign different values to Option A to ensure the variety. Option A's View count is constructed by having low (100-500) and high magnitude (1 million to 2 million) with two different levels of $\frac{Thumbs\ Up}{View}$ proportion (10%, 90%).

Figure 20 Option A's Choices in Study 2



Thus, with these four versions of Option A's View and the six corresponding Option B for each given Option A, we have 24 pairs of choice sets in total. Participants were randomly assigned into one version of Option A and were told to assume that the two alternatives in each choice set were similar to all other aspects except for the View and Thumbs Up.

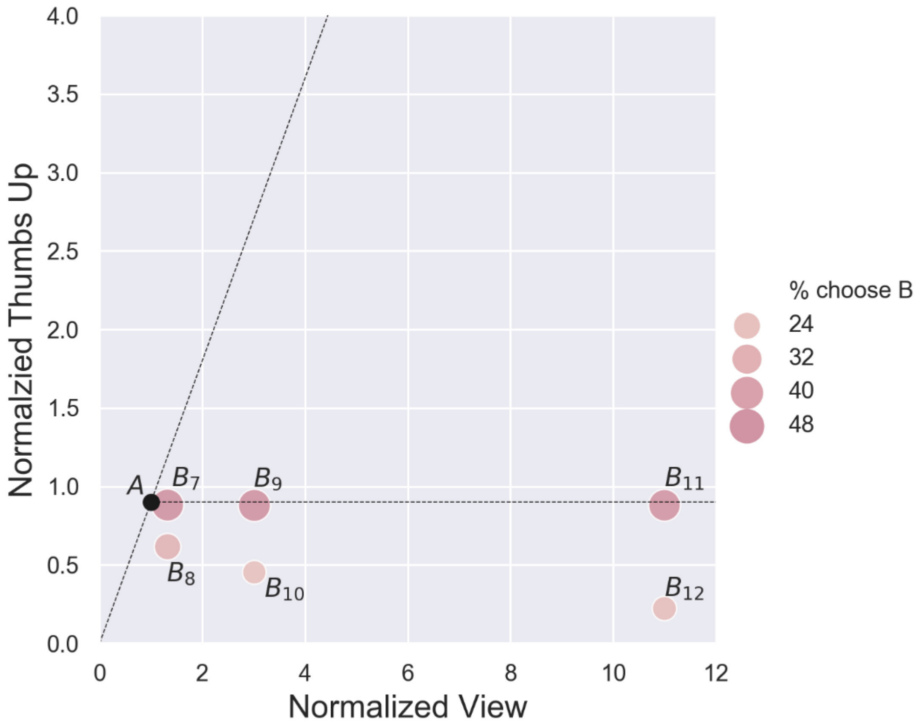
5.2. Results

Table 13 Results from Study 2

	B_7	B_8	B_9	B_{10}	B_{11}	B_{12}
Choose A	54%	69%	54%	75%	55%	74%
Choose B	46%	31%	46%	25%	45%	26%

Table 13 displays the percentage of participants choosing Option A and Option B. The results suggest the percentage of choosing Option B is significantly different in these six pairs (χ^2 = 18.16, p=0.003). While Option A is generally preferred over Option B in all six choice pairs, the percentage is different. When Option A has a similar Thumbs Up value as Option B (B_7, B_9, B_{11}), the percentage of participants preferring Option A over Option B is similar (54%, 54%, 55%). When Option B's Thumbs Up value is much smaller than Option A (B_8, B_{10}, B_{12}), the percentage of people preferring Option A over Option B increases. Figure 21 shows the scatterplot for our results.

Figure 21 Share of Option B in Study 1

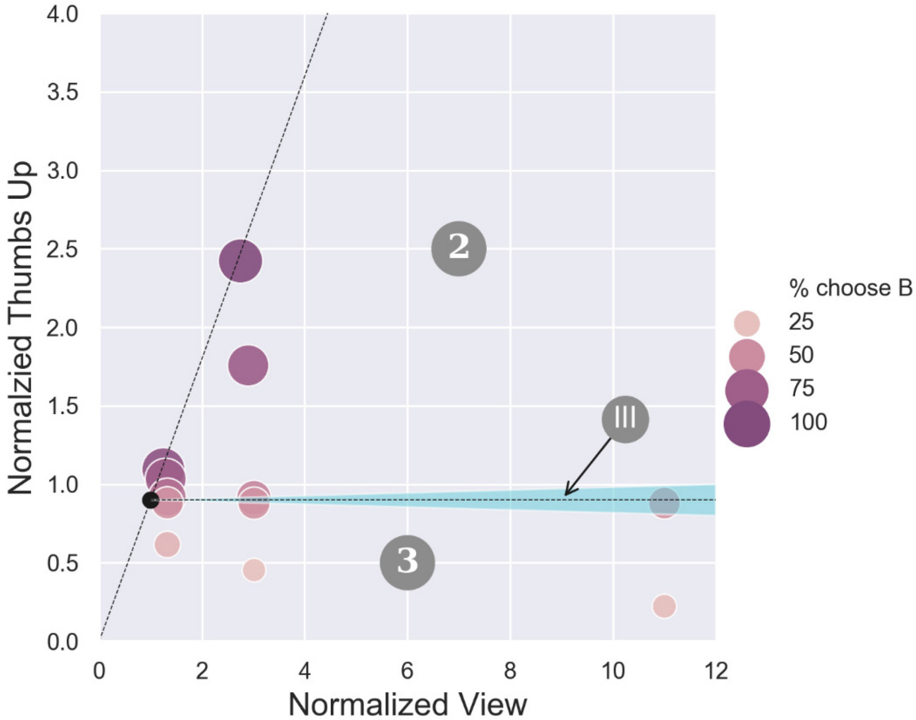


5.3. Discussion

Our results in the *conflicting region* suggest that consumers' decision-making heuristic is different from Bayesian Inference and three-stage decision process. Bayesian Inference predicts that all the consumers will choose Option A in this region. While the three-stage decision process suggests that when one option has a decisive advantage on one attribute, it will be more preferred. According to this decision strategy, consumers will choose Option B at choice pair B_{11}, B_{12}. But that's not what our results suggest. To better understand the underlying mechanism of how consumers behave, we plot study 1 and study 2's results together in Figure 22. We see that as Option B's location moves from the *seemingly dominant region* to the *conflicting region*, the percentage of consumers choosing B decreases. At the "similar Thumbs Up region", more people will prefer Option B over Option A when they are in the *seemingly dominant region*

(59%), but this percentage flip when Option B is in the *conflicting region* (46%, $\chi^2 = 5.63$, $p=0.018$).

Figure 22 Share of Option B in Study 1 & 2



When Option B moves from above the "similar Thumbs Up region" (region 2-III) to below the "similar Thumbs Up region" (region 3-III), the difference is that Option B changes from dominating on both View and Thumbs Up to only dominate on View and Option A becomes more dominate on Thumbs Up. This flip should not have an influence on the consumers who use signals to make decisions because Option A always has a higher *post-consumption signal*. However, the flip would influence those people who rely on **face value** in making decisions and think Thumbs Up is more prominent. For these people, their choices will change from choosing Option B to choosing Option A because Option A now has a larger Thumbs Up in the conflicting region. For the type of consumers who use the **minimum cognitive load**, the best decision strategy is to use face value when Option B is in Region 3-III because View is similar for Option A and Option B. Thus, they will choose the option with larger View. When Option B moves from Region 3-III to other locations in this *conflicting region*, the effort to use face value for decision-making increases because the conflict between Option A and Option B becomes larger. It would be more costly to use simply the face value for making decisions as it is hard to tell which option is better without considering the *post-consumption signal*, $\frac{Thumbs\ Up}{View}$ proportion. Therefore, the **minimum cognitive load** consumers will change their decision strategy from using face value to using signals. As a result, they will prefer Option A in the *conflicting region*. Table 14 summarizes the decision heuristics for the *conflicting region*.

Table 14 Consumer Decision Heuristics in Conflicting Region

	Cognitive Load	Region 3-III	Region 3	Decision Heuristic
Face Value		B	B	Choose the option with a larger View if View is more important.
		A	A	Choose the option with larger Thumbs Up if Thumbs Up is more important.
Minimum Cognitive Load		B	A	Use face value to make the decisions until the conflicts between options are significant. Then, choose the option with a larger *post-consumption signal*.
Signals	Cognitively bounded	A	A	Choose the option with a larger *post-consumption signal* when cognitively capable of detecting the superior option.
	Cognitively capable	A	A	Choose the option with a larger *pre-consumption signal* when the *post-consumption signal* is similar. Otherwise, choose the option with a larger *post-consumption signal*.
		A	A	Choose the option with a larger *post-consumption signal*.

6. Study 3: Presenting Post-consumption Signals

We have examined how different categories of consumers make decisions in Study 1 and Study 2. For consumers who want to rely on signals to make their decisions, some of them are cognitively bounded in using signals when Option B is located in the *seemingly dominant region*. Therefore, the main objective of Study 3 is to test how consumers will make decisions if the *post-consumption signals* are presented to them explicitly, which allows consumers to process this signal more easily.

6.1. Method

One hundred and twenty participants (50% male, M_{age}=41 years) were recruited through Amazon Mechanical Turk and were asked to participate in a survey on "understanding decisions". The experiment procedure is similar to the previous two studies. Participants completed six video choices and a basic demographic questionnaire. For each video, aside from the View and Thumbs Up information, we also presented the $\frac{Thumbs\ Up}{View}$ proportion.

Figure 23 Example of Choice Sets in Study 3

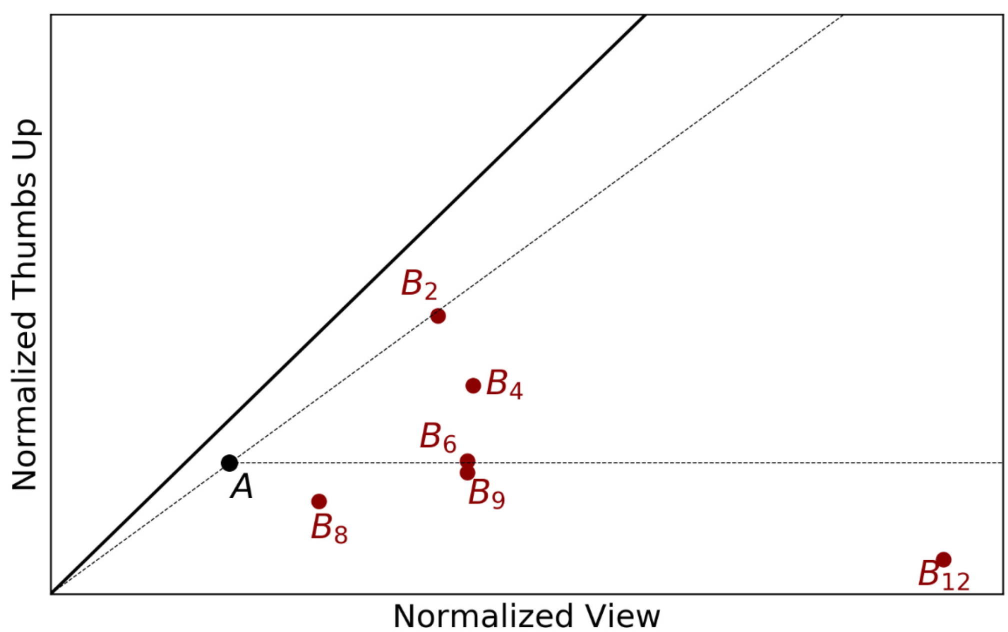
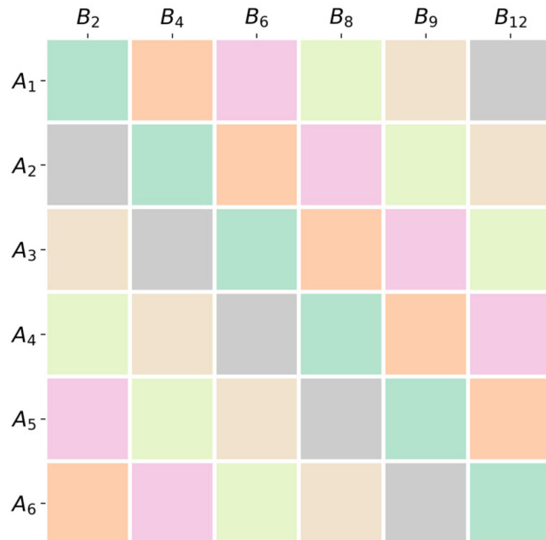


Figure 24 Randomization Strategy for Study 3



In order to compare the results with Study 1 and Study 2, where *post-consumption signals* are not explicitly presented, we choose the choice pairs that we have already tested in previous experiments. Figure 23 shows the six choice pairs we have chosen. Similar to Study 1, we also assign different values to Option A to ensure the variety. Option A's View count is constructed by having low (100-500) and high magnitude (1 million to 2 million) with three different levels of $\frac{Thumbs\ Up}{View}$ proportion (10%, 50%, 90%). Thus, with these six versions of Option A's View and the six corresponding Option B for each given Option A, we have 36 pairs of choice sets in total. Unlike Study 1 and Study 2, where participants were randomly assigned to one version of Option A, our randomization strategy is different here. If we still assign participants base on Option A's version, then it is very likely for participants to notice that one option's $\frac{Thumbs\ Up}{View}$ proportion does not change. In order to avoid this, respondents were randomly assigned six choice pairs from the total thirty-six pairs, with a balance between Option A's version and Option B' choices (Figure 24). Each color represents the six choice pairs one participant completes. Participants were randomly assigned to one of these color conditions. Participants were told to assume that the two alternatives in each choice set were similar to all other aspects except for the View and Thumbs Up.

6.2. Results

Table 15 Results from Study 3

	B_2	B_4	B_6	B_9	B_8	B_{12}
% Choosing B in Study 3	77%	30%	21%	16%	17%	13%
% Choosing B in Study 1 or 2	88%	76%	53%	46%	31%	26%

Figure 25 Share of Option B in Study 3

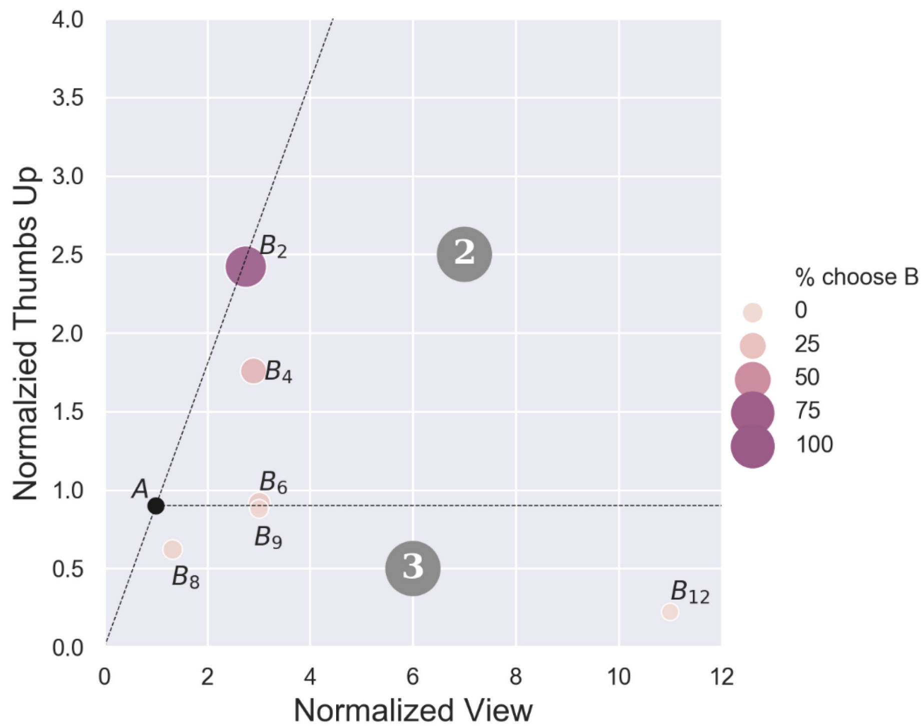


Table 15 presents the results from our experiments along with the results from Study 1 and Study 2 on the same choice sets. We can see that the percentage of consumers favoring Option B decreases when Option B moves from the top of the *seemingly dominant region* to the bottom of the *conflicting region* (Figure 25). The percentage of consumers choosing Option B decreases significantly from B_2 to B_4 (χ^2 = 59.12, p<0.001), and shows a decreasing trend from B_6 to B_{12}. The percentage of consumers choosing Option B is also smaller when consumers are explicitly presented with the *post-consumption signal*. Specifically, when Option B is at the "similar $\frac{Thumbs\ Up}{View}$ proportion region" (77% vs. 88%, χ^2 = 4.70, p=0.03).

6.3. Discussion

The decreased percentage of consumers choosing Option B from B_2 to B_4 clearly suggests the existence of cognitive bounded consumers who want to rely on the *post-consumption signal* for making decisions. When we explicitly present the information, these consumers are able to flip their choices and choose Option A. For the **minimum cognitive load** consumers, once the *post-consumption signal* is presented as face value, their behavior becomes similar to the **face value** consumer. We argue that as Option B's *post-consumption signal* becomes weaker and weaker compared with Option A, more consumers will prefer Option A, suggesting, the downward trend from B_6 to B_{12}.

Our results from Study 3 together with Study 1 and 2 confirms the hypotheses that there exist some consumers who are cognitively bounded in inferring the $\frac{Thumbs\ Up}{View}$ proportion. When $\frac{Thumbs\ Up}{View}$ proportion is presented, consumers will use this signal and prefer the option that has a larger value on this *post-consumption signal*. Therefore, we argue that once the $\frac{Thumbs\ Up}{View}$ proportion becomes explicit, all the people who use the **signal** for decision making will favor Option A. Similar to our previous categorization, among the people who use signal, there exists a proportion of people who think the small difference on *post-consumption signal* is incomparable to the large difference on the *pre-consumption signal*. However, we are not able to tease out this proportion of people in the cognitive bounded category in Study 1. The reason being Option B will always be chosen no matter these people are indifferent to the difference on *post-consumption signal* or they are cognitively bounded to tell the difference between Option A and Option B. Once the $\frac{Thumbs\ Up}{View}$ proportion is presented, consumers who are cognitively bounded and do care about the small difference in $\frac{Thumbs\ Up}{View}$ proportion will be able to choose Option A. Therefore, we see a decrease in percentage choosing B at B_2 between studies.

7. Conclusion

In this article, we examine how consumers make decisions when presented with multiple online social signals. Our experimental findings suggest different intuition compared with Bayesian Inference that predicts rational decision making and three-stage decision strategies that account for bounded rational consumers. We show that consumer's heuristic in making decisions depend on whether they simply use face value to make the decisions, or they consider more nuisance information, which is the online social signals. Consumers who only consider face values will prefer the product that has larger values in the prominent dimension. For consumers who want to use the nuisance signals for decision making, we categorize them into people who are cognitively capable and who are cognitively bounded. The cognitively capable consumers will rely on the *post-consumption signal* in making decisions. However, the cognitively bounded consumers will only use the *post-consumption signal* when it is cognitively easy to calculate. By explicitly providing the *post-consumption signal*, we find a sharp difference in consumer's choices that can only be attributed to these cognitively bounded consumers. Our study contributes to consumer decision making literature by proposing different types of decision heuristics. The decision heuristic we propose also confirms that consumers have a constructive choice process (Bettman et al. 1998). Their choices largely depend on the choice sets they experience. Our study also contributes to the current study on *pre-consumption signal* and *post-consumption signal* by proposing that these two signals impact on consumers jointly. Only examining one of these signals does not provide the whole picture in understanding how consumer make choices. We also contribute to the understanding of different online social signals by empirically show the *post-consumption signal* is stronger than *pre-consumption signal.* One practical implication for our study is that if platforms can display *post-consumption signals* explicitly, cognitive bounded consumers will benefit from this signal in making their decisions.

8. References

Balboni K (2020) 5 stars vs. thumbs up/down—which rating system is right for your app? *Appcues*. Retrieved https://www.appcues.com/blog/rating-system-ux-star-thumbs.

Banerjee A V. (1992) A Simple Model of Herd Behavior. *Q. J. Econ.* 107(3):797–817.

Bettman JR, Luce MF, Payne JW (1998) Constructive Consumer Choice Processes. *J. Consum. Res.* 25(3):187–217.

Bikhchandani S, Hirshleifer D, Welch I (1992) A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *J. Polit. Econ.* 100(5):992–1026.

Cai H, Chen Y, Fang H (2009) Observational learning: Evidence from a randomized natural field experiment. *Am. Econ. Rev.* 99(3):864–882.

Cheng X, Dale C, Liu J (2008) Statistics and social network of YouTube videos. *IEEE Int. Work. Qual. Serv. IWQoS*. 229–238.

Chevalier JA, Mayzlin D (2006) The Effect of Word of Mouth on Sales: Online Book Reviews. *J. Mark. Res.* 43(3):345–354.

Dewan S, Ho YJJ, Ramaprasad J (2017) Popularity or proximity: Characterizing the nature of social influence in an online music community. *Inf. Syst. Res.* 28(1):117–136.

Dewan S, Ramaprasad J (2012) Music blogging, online sampling, and the long tail. *Inf. Syst. Res.* 23(3 PART 2):1056–1067.

Dewan S, Ramaprasad J (2014) Social media, traditional media, and music sales. *MIS Q.* 38(1):101–121.

Duan W, Gu B, Whinston AB (2008) The dynamics of online word-of-mouth and product sales-An empirical investigation of the movie industry. *J. Retail.* 84(2):233–242.

Duan W, Gu B, Whinston AB (2009) Informational cascades and software adoption on the

Internet: An empirical investigation. *MIS Q.* 33(1):23–48.

Evangelidis I, Levav J (2013) Prominence versus dominance: How relationships between alternatives drive decision strategy and choice. *J. Mark. Res.* 50(6):753–766.

Fischer GW, Hawkins SA (1993) Strategy Compatibility, Scale Compatibility, and the Prominence Effect. *J. Exp. Psychol. Hum. Percept. Perform.* 19(3):580–597.

Fishbein M (1976) A Behavior Theory Approach to the Relations between Beliefs about an Object and the Attitude Toward the Object. *Springer*. 87–88.

Kahn BE, Meyer RJ (1991) Consumer Multiattribute Judgments under Attribute-Weight Uncertainty. *J. Consum. Res.* 17(4):508.

Keeney RL, Raiffa H (1979) Decisions with Multiple Objectives: Preferences and Value Trade-Offs. *IEEE Trans. Syst. Man Cybern.* 9(7):403.

Klein LR (1998) Evaluating the Potential of Interactive Media through a New Lens: Search versus Experience Goods. *J. Bus. Res.* 41(3):195–203.

Li X, Wu L (2018) Herding and Social Media Word-of-Mouth: Evidence From Groupon. *MIS Q.* 42(4):1331.

Liu Y (2006) Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue. *J. Mark.* 70(3):74–89.

Meyer RJ (1981) A Model of Multiattribute Judgments under Attribute Uncertainty and Informational Constraint. *J. Mark. Res.* 18(4):428–441.

Mudambi, Schuff (2010) Research Note: What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.com. *MIS Q.* 34(1):185.

Nelson P (1970) Information and Consumer Behavior. *J. Polit. Econ.* 78(2):311–329.

Nelson P (1974) Advertising as Information. *J. Polit. Econ.* 82(4):729–754.

Payne JW, Bettman R, Johnson EJ (1993) The adaptive decision maker: Contingencies in decision making.

Rosenberg MJ (1956) Cognitive structure and attitudinal affect. *J. Abnorm. Soc. Psychol.* 53(3):367–372.

Simon HA (1955) A Behavioral Model of Rational Choice. *Q. J. Econ.* 69(1):99.

Tucker C, Zhang J (2011) How Does Popularity Information Affect Choices? A Field Experiment. *Manage. Sci.* 57(5):828–842.

Tversky A, Kahneman D (1981) The framing of decisions and the psychology of choice. *Science (80-.).* 211(4481):453–458.

Tversky A, Sattath S, Slovic P (1988) Contingent Weighting in Judgment and Choice. *Psychol. Rev.* 95(3):371–384.

Tversky A, Simonson I (1993) Context-dependent preferences. *Manage. Sci.* 39(10):1179–1189.

Chapter 5. Conclusion

In response to the growing importance of online social signals in influencing consumer behavior, this dissertation set out to examine how various types of online social signals influence consumer behavior in different stages of the decision process. We first propose a conceptual framework to understand how different online signals can differently influence prospective consumer's different decisions (essay 1). By theoretically arguing for how these online social signals were generated by prior consumers, we show that online social signals have two dimensions: information type, and information source. The conceptual framework of online social signals on consumer decisions theoretically argues for how and why online social signals can have a differential impact on different stages. This study makes important contributions as it lays the theoretical foundation to explore online social signals. The conceptual framework also allows us to systematically review the findings from prior studies and identify the research gaps.

Essay 2 demonstrates how the popularity effect, taking into the influence of position, influence consumer's two-stage decision process in an online music context. The results from our experiments show while popularity influence consumer's search decision, this effect becomes smaller when we account for the position effect. However, neither popularity nor position has an effect on consumer choice decision conditioning on search. This study contributes to the literature by distinguishing the effect of popularity and position for different stages of decision making. It also disentangles the popularity effect from the position effect.

Essay 3 examines the influence of two online social signals, pre-consumption signals, and post-consumption signals, on consumers' choices of online videos. Through a set of experiments, we find that when the post-consumption signal requires an extra cognitive cost to infer, consumers will make their choices based on pre-consumption signals. When post-consumption signals require less effort to obtain, either through the conflicting option situation in which consumers are mandatory to spend effort or through the signal being explicitly presented, consumers will flip their decisions to choose the option that dominates on the stronger signal, which is the post-consumption signal. This study sheds light on the first essay by demonstrating that different social signals do have different strengths. And it also uncovers the underlying mechanism of how and why consumers make their choices between alternatives.

The findings from this dissertation not only contribute to the academic literature but also provides insightful guidelines for the practitioners. Platform designers can utilize the conceptual framework from essay 1 to display the most effective online social signals for helping consumers to make the decisions. The results from essay 2 suggest that platforms can leverage the position effect by charging a fee to display products on a higher position or featuring products on the top to reduce inequality among products. The findings from essay 3 imply that platforms can present post-consumption signals explicitly to consumers for them to make choices. However, if the goal of the platform is to increase popularity for the products, they should only show the pre-consumption signals.

Overall, the findings from these three essays in the dissertation could be the steppingstones for future research on online signals.

Bibliography

Adomavicius G, Bockstedt JC, Curley SP, Zhang J (2013) Do Recommender Systems Manipulate Consumer Preferences? A Study of Anchoring Effects. *Inf. Syst. Res.* 24(4):956–975.

Cai H, Chen Y, Fang H (2009) Observational learning: Evidence from a randomized natural field experiment. *Am. Econ. Rev.* 99(3):864–882.

Chevalier JA, Mayzlin D (2006) The Effect of Word of Mouth on Sales: Online Book Reviews. *J. Mark. Res.* 43(3):345–354.

Chintagunta PK, Gopinath S, Venkataraman S (2010) The Effects of Online User Reviews on Movie Box Office Performance: Accounting for Sequential Rollout and Aggregation Across Local Markets. *Mark. Sci.* 29(5):944–957.

Dayton E (2020) 10 Fascinating Amazon Statistics Sellers Need To Know in 2020. *Big Commer.* Retrieved (June 18, 2020), https://www.bigcommerce.com/blog/amazon-statistics/#a-shopping-experience-beyond-compare.

Dewan S, Ho YJ, Ramaprasad J (2017) Popularity or proximity: Characterizing the nature of social influence in an online music community. *Inf. Syst. Res.* 28(1):117–136.

Dewan S, Ramaprasad J (2012) Music blogging, online sampling, and the long tail. *Inf. Syst. Res.* 23(3 PART 2):1056–1067.

Evangelidis I, Levav J (2013) Prominence versus dominance: How relationships between alternatives drive decision strategy and choice. *J. Mark. Res.* 50(6):753–766.

Goes PB, Lin M, Au Yeung C man, Yeung C man A (2014) “Popularity effect” in user-generated content: Evidence from online product reviews. *Inf. Syst. Res.* 25(2):222–238.

Lee YJ, Hosanagar K, Tan Y (2015) Do I Follow My Friends or the Crowd? Information

Cascades in Online Movie Ratings. *Manage. Sci.* 61(9):2241–2258.

Marr B (2018) How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read. *Forbes Media LLC.*:19–22.

Mohsin M (2020) 10 Youtube Statistics That You Need to Know. *Oberlo*. Retrieved (June 18, 2020), https://www.oberlo.com/blog/youtube-statistics.

Mudambi, Schuff (2010) Research Note: What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.com. *MIS Q.* 34(1):185.

Salganik MJ, Dodds PS, Watts DJ (2006) Experimental study of inequality and unpredictability in an artificial cultural market. *Science (80-.).* 311(5762):854–856.

Simonson I (1989) Choice Based on Reasons: The Case of Attraction and Compromise Effects. *J. Consum. Res.* 16(2):158.

Spotify (2020) Spotify — Company Info. *Spotify*. https://newsroom.spotify.com/company-info/.

Tucker C, Zhang J (2011) How Does Popularity Information Affect Choices? A Field Experiment. *Manage. Sci.* 57(5):828–842.

Tversky A, Simonson I (1993) Context-dependent preferences. *Manage. Sci.* 39(10):1179–1189.

Zhang J (2010) The Sound of Silence: Observational Learning in the U.S. Kidney Market. *Mark. Sci.* 29(2):315–335.

Zhang J, Liu P (2012) Rational Herding in Microloan Markets. *Manage. Sci.* 58(5):892–912.